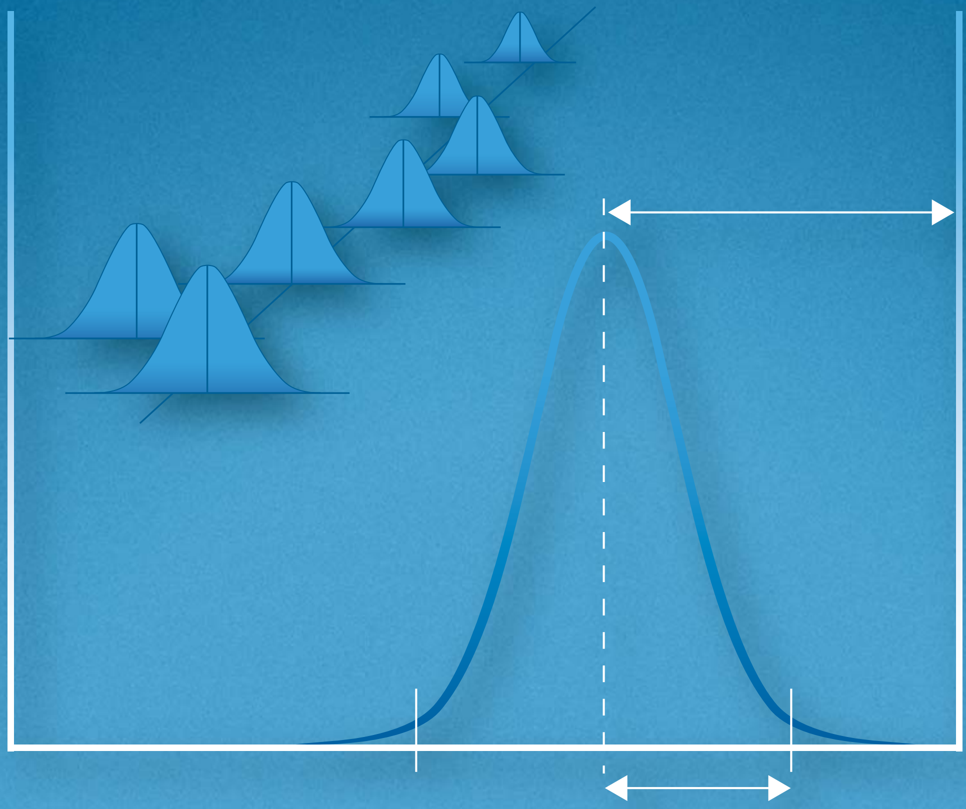


Quality Management in the Bosch Group | Technical Statistics

9. Machine and Process Capability

BOSCH
Invented for life

Quality Management in the Bosch Group

Technical Statistics

Booklet No. 9

Machine and Process Capability

Edition 11.2019

5. Edition 05.11.2019

4. Edition 01.06.2016
3. Edition 01.07.2004
2. Edition 29.07.1991
1. Edition 11.04.1990

All minimum requirements specified in this booklet for capability and performance criteria correspond to the status at the time of printing (issue date). CDQ 0301 is relevant for the current definition.

Table of contents

1 Introduction

The necessity of capability and performance records are obtained inter alia from the requirements of the standard [ISO 9001]. Some relevant excerpts in this regard:

> *The organization shall carry out the production and service providing under controlled conditions. If applicable, controlled conditions must include:*
>
> - *The implementation of monitoring and measuring activities at appropriate phases in order to verify that the criteria for controlling processes or results and the acceptance criteria have been met for products and services [ISO 9001; 8.5.1 c)];*
> - *The validation and regularly repeated validation of the capability to achieve planned results of the processes of the production or of service providing, when the resulting output could not be verified by subsequent monitoring or measurement [ISO 9001; 8.5.1 f)].*
> - *When a non-conformity occurs, including those arising from complaints, the organization has to respond and, if applicable, has to establish measures for monitoring and for corrections [ISO 9001; 10.2.1 a) 1)].*

These requirements were already included textually in previous versions and industry-specific manifestations of this standard. In simple terms, this means that the capability and performance records are especially necessary if product properties are not fully verified as a result of a manufacturing process, if for example a 100 % inspection can not be performed due to efficiency or cost reasons.

> *NOTE: A 100 % inspection means, that a particular characteristic is checked on every single manufactured part. Contrary to occasional misinterpretation, a 100 % inspection does not mean, however, that all available characteristics on a single part are checked.*

Basis of this booklet is the series of standard ISO 22514, and in particular the standard [ISO 22514-2]. Terms used are taken mainly from [ISO 22514-1], [ISO 3534-2], [ISO 3534-1], [ISO 9000] and [VIM]. Section “Terms and definitions” contains a selection of important definitions.

Purpose of a process is[1], to manufacture a product or to provide a service that meets predetermined specifications. The specifications are defined for one or more characteristics of the product or service. Capability and performance of a process can only examined and possibly detected in respect of a particular characteristic. This means that normally a separate proof must be provided for each characteristic. The characteristic may have a measurable physical quantity (e. g., length, current, temperature) or a countable property (e. g. within or outside of the tolerance interval).

The three essential steps to prove[1] process capability:

1. Assessing whether there is a stationary process which acts stable and therefore predictable over a reasonable period of time;
2. Determination of a statistical process model, i. e. a probability distribution for the process results with corresponding estimates for the parameters of this distribution;
3. Assessment of the process results on the basis of this process model, whether the properties of the produced product characteristics meet the specifications. If this is not the case, the process characteristics must be optimized in order to realize process results with the required properties.

Then make sure that the process characteristics, and thus the process, are not or only in predictable ways changing. This can be followed up e. g. with statistical process control [Booklet 7].

[1] Based on [ISO 22514-1; 1]

The steps are treated in this booklet from the viewpoints of *machine capability* as well as *short and long-term capability and performance of manufacturing processes*. The booklet is divided into the main part and the appendix. The appendix contains numerous references, additions and — if required on the basis of recurrent requests — explanations on theoretical foundations, which provide an increased demand on the mathematical understanding and preferably appeal readers with appropriate information needs.

It is expressly pointed out, that there is a great deal of corresponding "personal competence" required, when applying the contents and methods of this booklet. Meaningful Evidence of capability can never be the result of indiscriminately conducted evaluation of nonspecifically captured measurement data. The main objective of Evidence of capability with the required analysis is next to the determination of statistical indices, which meet the predetermined criteria, also to achieve a basic understanding of the process behavior.

2 Area of application

The statistical procedures and methods based on [ISO 22514-3] (capability of machines for manufacturing) and [ISO 22514-2] (capability and performance of manufacturing processes) described in Chapters 3 to 11 relate to one-dimensional, continuous quality characteristics. These approaches are applicable on processes in all industrial and economic sectors[2].

Section 14 leads in addition to the subject of two-dimensional, continuous characteristics [ISO 22514-6]. These include, for example, so called positional tolerances.

Appendix J finally illustrates an approach for the treatment of discrete characteristics based on [ISO 22514-5].

Typical applications[3]:

- There is the need to determine, if the processing result of a production machine or the result of a manufacturing process (incl. assembly) comply to the prescribed criteria and is acceptable, because a 100 % inspection is not performed (for example, for reasons of efficiency and/or cost reasons) or is not possible in principle (for example, in destructive tests).
- The Evidence of capability is a necessary prerequisite to release newly-starting or modified processes for series production (requirement based on internal rules and regulations) or required by the customer (e. g. due to contractual agreements).
- Determination of parameters for process control.
- Analysis and evaluation of results within the framework of the problem solving (for example, during an unexpected process behavior or in the cause analysis of field failures).
- Determination of measures for process optimization.
- Requirement for suppliers in order to be assured that required product specifications are met.

It is necessary to note that all values for parameters identified in the capability analysis represent only estimates of the true values i. e. a kind of "snapshot" of the current situation. It is therefore strongly recommended to determine and document the confidence intervals for these characteristics. Statistical software (for example, qs-STAT®) is usually already preset accordingly.

[2] Based on [ISO 22514-2; 1]

[3] Based on [ISO 22514-1; 1]

3 Flowchart

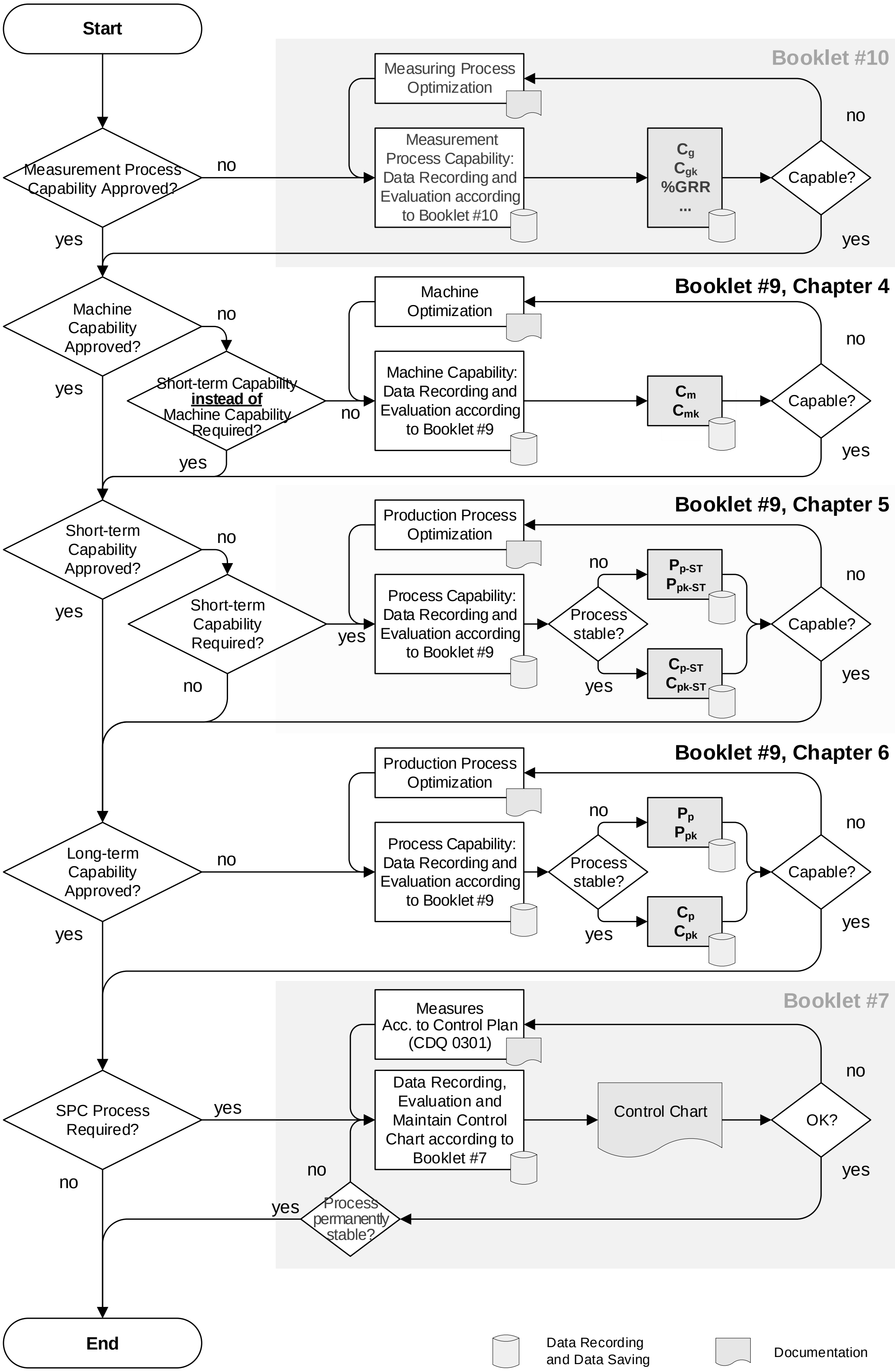


Figure 1: Operational sequence of capability analyses and process monitoring

4 Machine capability

The study of the machine capability is a short-term study. The aim is to detect and evaluate exclusively machine-related influences on the manufacturing process — and possibly understand.

However, numerous other influencing- and interference variables affect in addition. Typical examples of non-machine-related influences:

Man	• Personnel • Shift change	• ...
Machine[4]	• Rotational speed • Feed • Tools • Cycle times • Coolant flow and temperature • Pressure	• Current (for example on welding machines) • Performance (for example during laser welding) • Change status during optimization measures • ...
Material	• Semi-finished products and raw material from different deliveries or from different manufacturers • ...	• ...
Method	• Warm-up time of the processing device prior to random sampling • ...	• Different pre-processing, different production flow • ...
Milieu *(environment, „mother nature")*	• Room temperature *(e. g., change in temperature during the production of the random sample)* • Humidity, barometric pressur • Vibrations acting on the processing device	• Location of the processing device in the building (e. g. floor*)* • Extraordinary events • *(e. g. Open windows, heater on or off)* • *...*

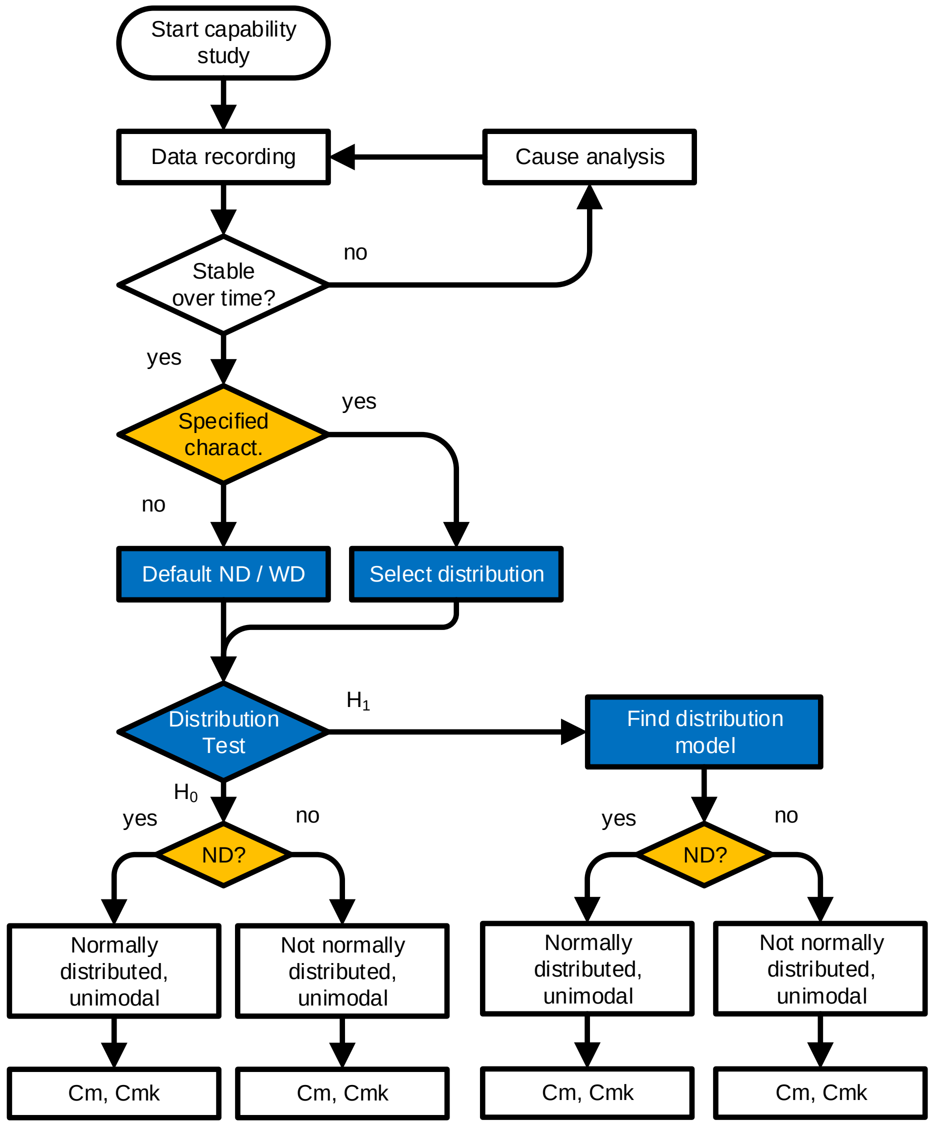
To exclude or at least minimize effects of non-machine related influences and disturbances, attempts are made to keep these variables as constant as possible. It is expected, that only influences from the machine and its changes affect the product and its characteristics.

If it is not possible in individual cases to keep non-machine-related influences constant (for example, room temperature), the changes or the respective influences have to be recorded and documented. This information may provide approaches for optimization measures if capability criteria are not met.

Requirement for a machine capability study are capable measurement and test processes (see Section 9.3 in this Booklet 9, [Booklet 10] and [CDQ 0301]).

[4] The listed factors are examples of parameters whose settings are not determined by the machine. Machine related however, are variations and deviations from these settings, which usually cannot be influenced or controlled by the user.

Procedure and evaluation configuration



See Section

4.1

4.2

4.3

Appendices C and D

4.4, 4.5

Figure 2: Flowchart depicting the procedure of a machine capability study

4.1 Data collection

The data collection for a machine capability study is carried out in three steps:

1. Preparation of the processing device, so that the measurements are as much as possible in the middle of the tolerance interval (pre-run).

   For unliteral limited characteristics the best possible setting regarding the target value should be selected.
2. Preparation of a representative number of pieces (at least $n = 50$, preferably $n = 100$) in a continuous production run, i. e. in an uninterrupted sequence. Deviations have to be documented.
3. Measurement of the part characteristic or the part characteristics and documentation of the results according to the production sequence.

It is essential for the analysis that all n measurements are considered as a single sample ($m = 1$) with sample size n.

4.2 Investigation of the temporal stability

First of all, it is qualitatively assessed based on the original value chart whether the registered individual measurements are stable over time or not:

- Are there systematic changes over the course of time?
- Do the individual values concentrate in the vicinity of the established set value?
- Are all of the individual values within a zone that corresponds to about 60 % of the tolerance interval?

Evidence that the process is not stable is revealed, for example, by the following observations:

- There are some unexplained outliers (response of the outlier test or values outside the plausibility limits).
- There are inexplicable jumps, steps or trends.
- Individual values are mostly above or below the setpoint.
- In a bilateral limited characteristic are the individual values mostly in the vicinity of the limit values.

If the value curve is not plausible, the causes for this behavior are to be investigated and eliminated. Subsequently, the capability study must be repeated.

4.3 Investigation of the statistic distribution

The knowledge of the production process and the type of tolerance specification often allow a conclusion as to the distribution model, which is suitable to describing the empirical distribution of the characteristic.

- One can, for example, expect, that a characteristic, whose value differs equally likely both upwards and downwards from the setpoint (positive or negative deviation), is approximately normally distributed. But this is not necessarily always the case.
- In contrast, characteristics, which possess a unilateral "natural" limit, follow generally a skewed, asymmetric distribution. For example, concentricity and roughness can never assume negative values by definition. The value 0 in this case is a natural lower limit ($LL^* = 0$). Experience shows that many zero-limited characteristics of shape and position according to [ISO 1101] can be described with an amount distribution of type 1 or 2.

Within the scope of a machine capability study, a distinction is therefore made according to the following scheme. The concrete distribution allocation is then carried out with the aid of the procedures according to Appendix C, in particular C.2.

Characteristic (measurand) specified	• Normal distribution • Folded normal distribution • Rayleigh distribution
Characteristic (measurand) not specified	• bilaterally limited: Normal distribution • unilaterally limited: Weibull distribution

It is expressly pointed out that a characteristic *may* behave in accordance with these rules, but *does not have to* do so in principle. In individual cases, clear deviations can be observed (see also notes in Section 6.5).

For this reason, the next step is to perform a test for a specified distribution. The diamond (rhombus; decision symbol) contains both a chi-squared test and several tests for normal distribution, which are selected according to the sample size (see Appendix D.7).

If the predefined distribution model was rejected by the test for normal distribution or the chi-squared test, a distribution matching the values is determined.

Ultimately, it is up to the user to check and assess whether a meaningful distribution model has been assigned to the available measurement data in consideration of all statistical circumstances and technical conditions.

4.4 Calculation of capability indices

Basics and methods of calculating capability indices are shown in chap. 8.

As **standard method** for calculating capability indices, the quantile method $M_{2,1}$ is used according to [ISO 22514-2] (see Section 8.2). This is the only method that can be used for all distribution models without limitation. However, the calculations require the use of appropriate statistical software (for example, qs-STAT®).

Data that can be described in good approximation by a **normal distribution** could be evaluated alternatively with all methods possible according to [ISO 22514-2] (see Section 8.3 and Appendix E, Table 4). The number of samples has to be set to $m = 1$. In practice, usually the method $M_{3,5}$ is used[5]. These calculations can also be performed “manually” (or e. g., using MS-EXCEL®).

Other unimodal distributions could be evaluated, alternatively to $M_{2,1}$, especially with method $M_{4,5}$ according to [ISO 22514-2] and $m = 1$, which also allows “manual” calculations (see chap. 8.3). Since no information on the distribution model is present and therefore s_{total} is used, this method results in comparison to $M_{2,1}$ normally to the most disadvantageous results.
For all other distribution models, with few exceptions (see Appendix E, Table 4) is the quantile method $M_{2,1}$ required.

4.5 Criteria for machine capability

The indices for machine capabilities are denoted by C_m and C_{mk}.

NOTE 1: Differing designations are possible according to customer requirements.

NOTE 2: [ISO 22514-3] uses the variable names Pm and Pmk instead of Cm or Cmk.

Decisive for the current limits is [CDQ 0301] in the current version. At the time of the publication of this edition of Booklet 9 the following requirements and limits apply:

	Requirement
Number of parts (measurements)	$n \geq 50$ *($n \geq 100$ recommended)*
Potential capability index	$C_m \geq 1.67$
Critical capability index	$C_{mk} \geq 1.67$

NOTE 3: The limits are to be regarded as absolute minimum requirements, which should not be undershot. The requirements may be increased depending on the application.

If the capability criteria are not met, a cause analysis and possible repetition of the capability study is needed.

[5] In the case m = 1, some calculation methods according to [ISO 22514-2] are identical (for example, M3,2 and M3,5, M2,5 and M4,5)

4.6 Machine capability with reduced effort

According to Section 4.5, the required number of manufactured and to be measured parts is at least $n = 50$ (better $n = 100$). In practice, however, it may be unavoidable in **exceptional cases** to manage, regardless of this setting, a reduced sample size $n < 50$, for example, with very elaborate measurements associated with exceptionally high costs, or destructive tests.

However, with decreasing sample size, the reliability of the information decreases because the confidence interval of the calculated characteristic value becomes larger. This can be compensated to a certain degree by increasing the minimum level of capability indices (see Appendix I.1).

Statistical software is now often pre-set (for example, qs-STAT®) so that this increase is automatically set depending on the sample size.

NOTE 1: The basic idea is, that the lower confidence limit, which applies for $n = 50$ parts, can not be by undershot even in the case of $n < 50$ parts. Further details are explained in Appendix I.1.

NOTE 2: Reductions of a sample size below $n = 50$ should always be coordinated with the local, quality responsible entity.

Unless appropriate statistical software is available, the following two-step procedure is alternatively possible:

1. Of the $n = 50$ consecutively manufactured parts initially only every 2nd part is measured, i. e. the parts no. 2, 4, 6, ..., 50. With this, 25 readings are obtained (per characteristic). The machine is capable, if the calculated capability index from these 25 values meets the criterion $C_{mk} \geq 2{,}0$.
2. If only $1{,}67 \leq C_{mk} < 2{,}0$ is reached, the remaining 25 parts no. 1, 3, 5, ..., 49 are measured and the already present measurements are supplemented by these measurement results. The machine is capable, if the calculated capability index from these 50 values meets the criterion $C_{mk} \geq 1{,}67$.

The first step of this procedure is similar to a reduction of the sample size to $n = 25$ parts and a corresponding increase of the minimum requirements. The second step involves a kind of "additional option to repair" if the target is not reached with step 1. However, it is only in exceptional cases plausible (for example, with very elaborate measurements), if not all the parts are measured and evaluated that are available anyway.

NOTE 3: The contractually stipulated requirement with regard to the machine manufacturer continues to remain $C_{mk} \geq 1{,}67$ at $n = 50$ parts. A machine acceptance in accordance with the above procedure ($C_{mk} \geq 2{,}0$ at n = 25 parts) is however also allowed at the machine manufacturer.

5 Process capability (short-term)

In studies of machine capability (Ch. 4), characteristics of product parts are evaluated which have been manufactured in a continuous production run in an uninterrupted sequence, so that **possibly only the** influence of the machine is active.

In contrast, the parts to be measured originate in studies of long-term process capability (Chapter 6) from a larger, more representative period for the series production, so that **possibly all** influences on the process, which are to be expected, take effect.

Particularly during series rump-up, there are frequently not enough product parts available nor can enough parts be taken out of the manufacturing process over a sufficient period of time. Nevertheless, as an alternative, or in addition to the machine capability, at least a preliminary conclusion about the expected manufacturing process capability can be demanded (see *"Initial Process Capability"* [AIAG PPAP] and *"Preliminary process capability"* [VDA-4]). In this case, a short-term study is conducted, which can differ from the long-term study in the following points.

- **Type of sampling:** The parts to be studied can be taken out of the manufacturing process in shorter intervals, if necessary in extreme cases, one immediately after the other.
- **Number of parts:** It is permissible if there are not sufficient parts available, to take less than the required 125 parts for the long-term study.
- **Limits for capability and performance indices:** The increased limit of 1.67 applies if there are more than 125 parts. If there are less than 125 parts, the limit is raised depending on the number of parts at the same value as in the long-term study with reduced quantities (see Appendix I).
- **Designation of statistical indices:** Capability indexes are designated with C_{p-ST} und C_{pk-ST} and performance indexes with P_{p-ST} und P_{pk-ST} (**s**hort-**t**erm).

Evaluation of the data takes place in exactly the same manner as in the long-term study (Chapter 6).

NOTE: Information about random sample sizes and capability requirements are valid at the time of publication of this directive. Decisive is the currently valid edition of [CDQ 0301].

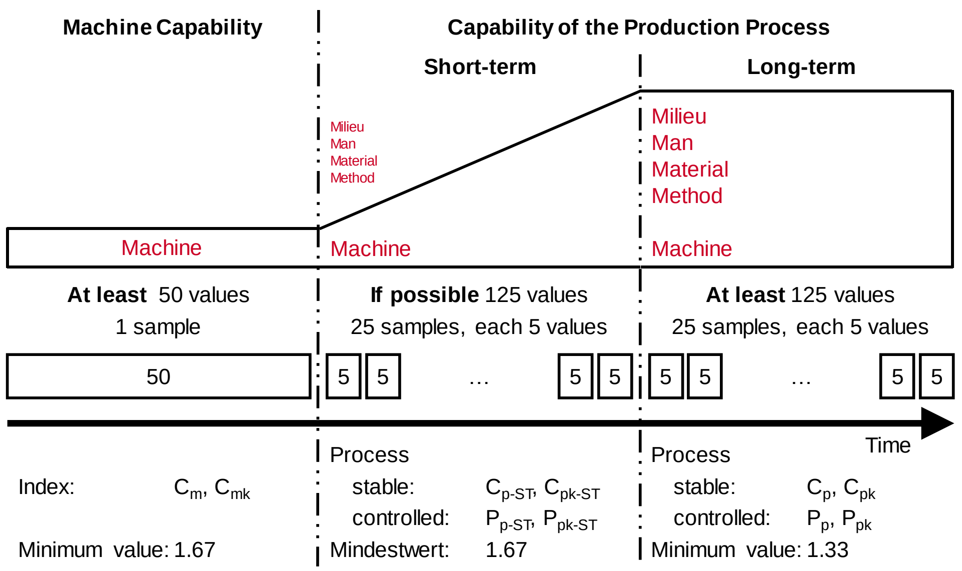


Figure 3: Machine capability, short and long-term capability of a manufacturing process

The less parts available and the shorter the sampling period, so that the parts must be taken from the manufacturing process in an increasingly immediate sequence, the more the short-term capability corresponds to the machine capability. The more parts available and the longer the period over which the parts are taken, the more the short-term capability corresponds to the long-term capability. Basically, the short-term study should come as close as possible to a long-term study.

6 Process capability (long-term)

The process capability is the result of a long-term study. In addition to the pure machine-related influences, all possible influences should be detected which affect the manufacturing process for a longer operating time. These disturbances can be summarized in categories by the superordinate concepts man, machine, material, method and milieu, often abbreviated by *5M*.

Evaluation configuration

The evaluation configuration for process capability essentially consists of the 3 areas (main branches), which are highlighted in grey in the following figure. The distribution models (resulting process distributions) according to [ISO 22514-2] of the left area are selected if the position and variation of the process are stable. If the process behavior is unstable, the evaluation algorithm branches to the middle or right area. The associated parameters are then called the performance indices P_p and P_{pk}.

Process behavior:	Process variation stable Process location stable	Process variation stable Process location not stable	Process variation instable
Distribution model:	A1, A2	C1, C2, C3, C4	B, D
Index:	C_p, C_{pk}	P_p, P_{pk}	P_p, P_{pk}

⇓ ⇓ ⇓

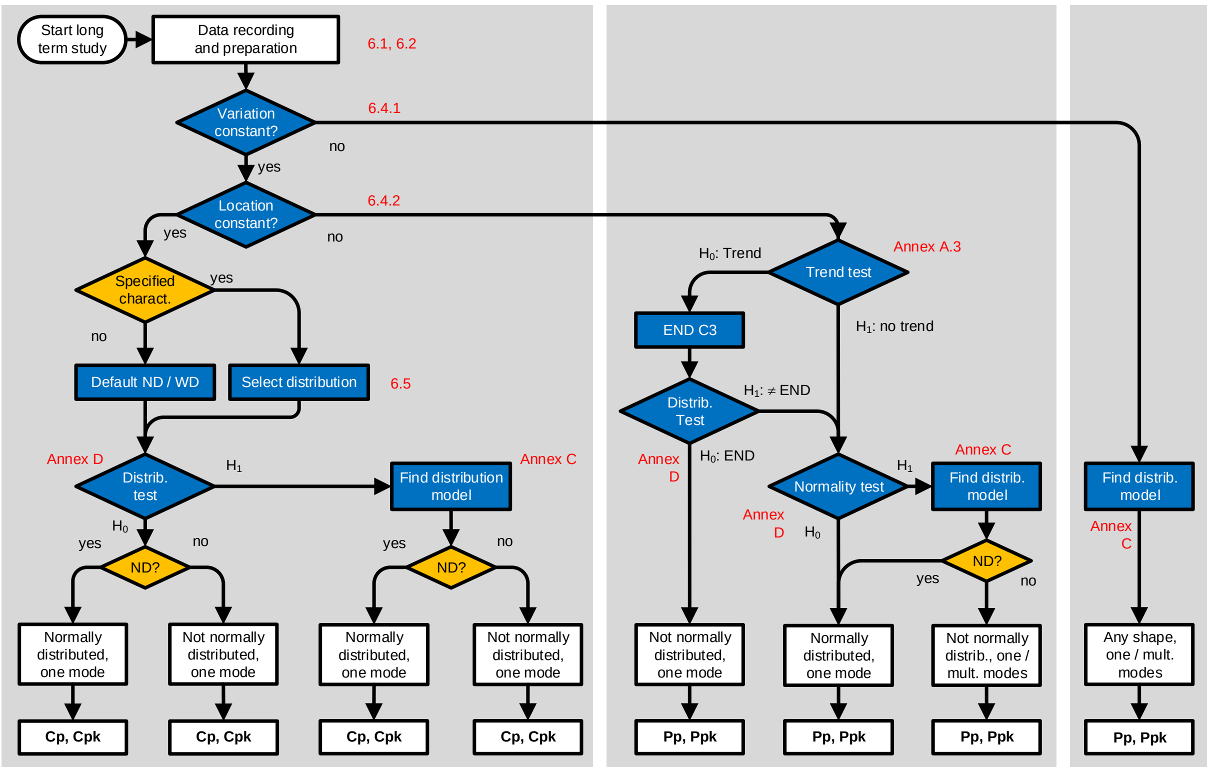


Figure 4: Procedure of a long-term study for manufacturing process capability

6.1 Data collection

The study requires a representative number of the production quantity, but at least 125 non-selected parts (for example, $m = 25$ samples each with sample size $n = 5$) over a sufficiently long period of time, so that all possible expected influences can have an effect on the process.

NOTE 1: In special cases, it may be imperative to use less than 125 parts, e. g. with very elaborate measurement methods or with destructive tests. In these cases, however, the "reliability" of statistical significance with decreasing sample size will be reduced, meaning, the confidence interval for the calculated characteristic value becomes larger. This can be compensated to some degree by raising the limits on capability and performance indeces (see Appendix I.1 for more information). Reduced sample sizes should be coordinated with the quality responsible entity.

NOTE 2: [VDA 4] recommends a minimum of 20 production days. This is, however, only be seen as a rough estimate. Instead, the time period should be appropriately aligned on the behavior and the boundary and ambient conditions of the respective process.

Three cases have to be distinguished for data collection.

Case 1: First evidence of capability during the initial phase

For the first analysis during the initial phase of a manufacturing process the data is usually collected specifically for the determination of capability indices.
At this time, usually no usable information on the process behavior exists, so that a tailored interval has to be estimated for this process of taking the individual samples. National and international standards and guidelines contain no guidelines or indications for that. The procedure described below is a common, semi-quantitative approach in practice.
If at maximum 75 % of the tolerance interval should be used (corresponding to $C_{pk} = 1.33$) and if the target value is in the middle of this interval, the deviations of the sample means of this setpoint must not be greater than ± 37.5 % of the tolerance T.

NOTE 3: At higher requirements such as max 60 % or 50 % tolerance utilization (corresponding to $C_{pk} = 1.67$ or $C_{pk} = 2.00$), the maximum acceptable deviations are reduced to + 30 % or + 25 % of tolerance T.

NOTE 4: If the specification is not in the middle of the tolerance interval, this asymmetry has to be taken into account. Example: Setpoint 5.0 mm, specification limits 4.9 mm and 5.4 mm, i. e., tolerance 0.5 mm, including 0.1 mm (20 %) below and 0.4 mm (80 %) above the setpoint; standardized limits at the highest 75 % tolerance utilization: -15% (corresponding to -0.075 mm, i. e., 75 % of -0.1 mm) or +60 % (corresponding to 0.3 mm, i. e. 75 % of 0.4 mm). Zero limited characteristics represent an extreme case of asymmetrical intervals, meaning that deviations in one direction are omitted.

For visualization of the process behavior, the displaying of the mean values as a time series graph with boundary lines at a distance of e. g. ± 37.5 % of the tolerance T to the desired value is appropriate.

NOTE 5: Such a time series chart is only a temporary aid and should not be mixed up with a control chart or used as such.

In general, it makes sense to start with the shortest possible sampling interval (For example, several samples per shift) and gradually adjust this interval to the observations.

- All mean values are within the deviation limits, the changes from value to value are clearly visible and unsystematic (randomly): Sample interval appropriate, no measures necessary.
- All mean values are within the deviation limits, however, no or only minimal changes from value to value are recognizable: Sample interval possibly too short; extend interval (e. g. double); repeat adaptation as necessary repeatedly.
- Some mean values are outside the deviation limits: Sample interval may be too long; shorten interval (e. g. halve); repeat adaptation as necessary.

If multiple adjustments of the sampling interval are not successful, look for possible causes for this process behavior and possibly eliminate these.

Case 2: Re-validation of the capability of processes, which are monitored by means of control charts

If a running production process is controlled using control charts (SPC process), the most recently collected data of the control charts are used for the regular revalidation of the process capability. The sampling interval which is used during a single revalidation, is therefore identical to the sampling interval of the control charts [Booklet 7].

Case 3: Re-validation of the capability of processes, which are not monitored by means of control charts

For all other processes, the data will be collected at specific time intervals as in the initial phase, specifically for checking and ongoing verification of the capability indices. For a single revalidation, typically the sample interval used is the one that was determined in the initial phase and was possibly further optimized later.

In all cases the characteristics are measured on each part and the results for each part are documented according to the production sequence. The evaluation of the measured data takes place regardless of how the data was collected.

6.2 Outliers

Data evaluation in the context of process capability analysis requires that the data set to be analyzed does not contain any “outliers”. This applies in particular to the statistical tests in Section 6.3 as well as in Appendices B and D. Outliers are individual values that “differ significantly from the majority of the other data”. Causes can be, for example, human errors (input and transmission errors, mix-ups, operating errors) or defect measurement systems.

Outliers can, for example, lead to

- the selection/allocation of a different distribution class (symmetrical, asymmetrical) or
- the selection/allocation of a different individual distribution or to
- a greater estimation of the variation range, or
- fake systematic temporal changes.

The evaluations of a data set with and without the potential outliers generally provide different results in the capability or performance indices.

It cannot be excluded with certainty that values initially regarded as outliers will subsequently turn out to be correct measured values and important information. If one or more outliers were excluded from the evaluation for small data sets, it should be observed whether and how the evaluation results change over time as the data basis increases.

But what is a "clear difference to the majority of data"? In order to objectify the decision, the following options are usually used.

Plausibility limits

qs-STAT® offers the possibility of setting so-called plausibility limits. For example, you can exclude technically impossible value ranges. If a characteristic value lies outside these limits, it is regarded as an outlier and excluded from the evaluation (but not deleted).

Outlier test

It is also possible to examine the data set for outliers on the basis of statistical criteria. However, many outlier tests listed in the literature are based on a concrete distribution model. In the present case, these tests cannot be used, since the allocation of a distribution model takes place later.

The Hampel outlier test considers the absolute deviation of a potential outlier from the median of the individual values in relation to the variation of the individual values [Hampel], [Sachs], [Schulze].

$|r_i|$ is the absolute deviation of the single value x_i from the median $\tilde{x}$ of the single values: $|r_i| = |x_i - \tilde{x}|$. $\tilde{r}$ is the median of these deviations $|r_i|$ and is also called "median absolute deviation" (MAD). x_i is considered an outlier if $\frac{|r_i|}{\tilde{r}} \cdot 0.6745 \geq T_{(n;\,1-\alpha)}$. For $n = 125$ and $(1-\alpha) = 95\ \%$ is T=3.8. The test is also able to detect several outliers.

NOTE: In the case of a normal distribution the quantity $\frac{\tilde{r}}{0.675}$ *is an unbiased estimator of the sample's standard deviation* σ.

In this original form described by Hampel, the test is intended for symmetrical distributions. Weaknesses in the application to asymmetric distributions can be eliminated by the approach of [Kölling]. In qs-STAT® different limit values are calculated for outliers below and above the median using Hampel's critical value $T_{(n;\,1-\alpha)}$:

$$HG_p = \tilde{x} - \frac{\tilde{x} - \hat{Q}_p}{u_{1-p}} \cdot T_{(n;\,1-\alpha)} \qquad HG_{1-p} = \tilde{x} + \frac{\hat{Q}_{1-p} - \tilde{x}}{u_{1-p}} \cdot T_{(n;\,1-\alpha)}$$

Example:

$$HG_{5\%} = \tilde{x} - \frac{\tilde{x} - \hat{Q}_{5\%}}{u_{95\%}} \cdot T_{(n;\,1-\alpha)} \qquad HG_{95\%} = \tilde{x} + \frac{\hat{Q}_{95\%} - \tilde{x}}{u_{95\%}} \cdot T_{(n;\,1-\alpha)}$$

$\hat{Q}_p$ is the empirical p% quantile of the sample, u_{1-p} the $(1-p)$ quantile of the standard normal distribution. For the parameter p [Kölling] recommends different values depending on the sample size n, e. g. $p = 5\ \%$ for $n = 125$.

6.3 Classification and rounding

Classification

Obviously the classification (grouping), i. e. the number of classes and the class width, has a big influence on the "appearance" of a histogram. In the statistical literature there are numerous rules of thumb for this choice or definition (see e. g. [Booklet 1]). Reasons for this are, for example, the comparability of different data sets of the same characteristic or the relevance in the application of the Chi-squared test. qs-STAT® selects an integer multiple of the resolution as the class width.

Rounding

Rounding means replacing the value of a number by the integer multiple of a place value [DIN EN 80000-1]. This topic concerns on the one hand the aspect of rounding rules (see [DIN 1333]), but on the other hand also the metrological aspect.
According to [GUM], numerical values for a measurement result x may not be given with an excessive number of digits. In particular, it does not make sense to indicate the measurement result with more than one additional decimal place than corresponds to the resolution of the measurement system. Further decimal places cannot be determined with the measuring equipment used and are therefore worthless (cf. [Booklet 8]).

Note: Rounding, grouping and classification can influence the allocation of the distribution model and thus the results for the capability and performance indices. For example, the classification of frequencies in the histogram can have an influence on the result of the Chi-squared test.

6.4 Investigation of the process stability

The next step is to determine whether the measurements are stable over time or not. The aim of the analysis is to be able to specify a characteristic value regardless of whether the process exhibits a trend or batch jumps or not.

Characteristics of stable processes are the following, content equivalent information:

- mean and variation are (nearly) constant.
- No systematic mean value changes occur (for example trends, batch jumps).
- There is no significant difference between sample variation and total variation.
- Each individual sample represents with regard to location and the total process variation.

The investigation of the time series initially covers the aspects of stability of the process variation and stability of the process location. If the situation is unstable, a test for trend is carried out for further differentiation. See Appendix A.3.

Test for stability of the process variation

An unstable process variation indicates that the process behavior basically is not statistically explainable and the process is therefore not in control. It is necessary to investigate the causes of this "chaotic" behavior, to eliminate it and repeat the capability study.

- Cochran's C-test can be used to determine whether the largest of the variances of k samples is significantly different from the variances of the remaining samples. See Appendix A.1.1.
- Analysis of variance and F-test (Appendix A.1.2 and [Booklet 3])
- The test in Appendix B.2 uses the variation of the individual samples.

Test for stability of the process location

An unstable process location indicates that there are non-random influences on the process behavior. In this case, the evaluation configuration branches into the middle range, see Fig. 4. The following options are available for stability testing:

- Kruskal-Wallis test (Appendix A.2.1)
- Location of the individual samples (Appendix B.1)
- Standard deviation of the sample means (Appendix B.3)

6.5 Investigation of the distribution

The measured characteristic values are interpreted as realizations of a statistical random variable.

In particular, it is assumed that the measured parts are representative of the parts to be manufactured in the future. When talking about a “population” in this context, one must not forget that this is a fictitious, not yet existing set of objects.

Formulations such as “Determining the distribution type” may possibly give the impression that behind the measurements is a specific distribution which is initially not known, but can be objectively determined with statistical methods.

In fact, you can only select one distribution model (for example, normal distribution) and verify by statistical tests whether the measurements with this model are at least approximately compatible (see Appendices C and D). All other derived statements from this model stand or fall with its validity.

Eight qualitative resulting process distributions are presented in the standard [ISO 22514-2], which are suitable for describing real production processes. “Qualitative” here means, that it is merely stated, how the resulting distribution arises from a “momentary distribution” with time-varying location and variation parameters and whether a single or multimodal distribution arises in the process.

> *NOTE 1: The “momentary distribution” can be understood as a distribution, which is represented by a single sample, whose individual values are recorded almost simultaneously, that is with very little time interval. The temporal evolution of the “moment distribution” is then represented by the distributions of the various individual samples, which are captured in greater time interval from one another.*
>
> *NOTE 2: Thus also the view of [AIAG SPC] becomes understandable, that determined on the basis of individual samples, capabilities are regarded as “short-term capabilities” (see also Appendix H).*

Therefore the task of choosing the *“correct”* distribution, meaning selecting the most appropriate distribution for available measurement data taking into account the technical conditions exists for the user.

> *NOTE 3: Choosing of appropriate distribution models must take into account, which models are in principle possible and appropriate, based on the known physical and technical conditions. The choice can not be done arbitrarily, that means, on purely mathematical basis from the viewpoint of achieving an optimal result for the capabilities (see Appendix C.4).*

The three tests according to Section 6.4 for stability of variation, stability of location, and trend represent a kind of “filter” with which a rough allocation to one of the resulting process distributions according to [ISO 22514-2] is made (see Appendix E, Figure 27). The following distributions can then be assigned explicitly in a suitable manner using the procedures according to Appendix C, in particular C.2:

- Normal distribution
- Logarithmic normal distribution (lognormal distribution)
- Folded normal distribution
- Rayleigh distribution
- Weibull distribution
- Extended normal distribution
- Mixture distribution

If there is no automated distribution assignment available, the selection of a best fitted distribution can for example be facilitated by showing the individual values in the probability plot of the distributions in question (see also Appendix C) or by statistical goodness-of-fit tests.

6.6 Calculation of capability and performance indices

The methods for calculating capability and performance indices are shown in Chapter 8.

As **standard method** for calculating of indices of process capability and performance indices the quantile approach $M_{2,1}$ is used according to [ISO 22514-2] (see Section 8.2). This is the only method that can be used without limitation for all resulting process distributions according to [ISO 22514-2]. However, the calculations require the use of suitable statistical software (e. g., qs-STAT®).

Data that can be described in good approximation by a **normal distribution with stable location**[6] could be alternatively evaluated with all possible methods according to [ISO 22514-2] (see Section 8.3 and Appendix E, Table 4). In practice, usually the methods $M_{3,2}$, $M_{3,3}$, $M_{3,4}$ and $M_{3,5}$ are used. All alternative calculations can also be performed "manually" (or e. g., using MS-EXCEL®).

Other unimodal distributions with stable location[7] could alternatively to $M_{2,1}$ still be evaluated in particular with methods $M_{2,5}$ and $M_{4,5}$ for according to [ISO 22514-2], which also allows "manual" calculations (see Section 8.3). Since no information on the distribution model is present and therefore s_{total} is used, result in these methods generally less favorable results in comparison to $M_{2,1}$.

For all remaining distribution models, the quantile method $M_{2,1}$ is required with few exceptions (see Appendix E, Table 4).

6.7 Criteria for process capability and performance

Long-term capability and performance of a production process require continuing compliance with predetermined limits. [CDQ 0301] in its currently valid version is applicable for the current limit value. At the time of the publication of this edition of Booklet 9 the following requirements and limits apply:

Stable process	Requirement	Process in control	Requirement
Number of parts (measurements)	$m \cdot n \geq 125$	Number of parts (measurements)	$m \cdot n \geq 125$
Potential capability index	$C_p \geq 1{,}33$	Potential performance index	$P_p \geq 1{,}33$
Critical capability index	$C_{pk} \geq 1{,}33$	Critical performance index	$P_{pk} \geq 1{,}33$

NOTE: The limits are to be regarded as absolute minimum requirements, which should not be exceeded and should be determined based on, for example, $m = 25$ samples with sample size $n = 5$ so that $m \cdot n \geq 125$ parts (measuring values). The requirements can be raised in dependence on the application case.

If the capability or performance criteria are not met, a cause analysis and possible repetition of the capability study is needed.

---

6 Distribution model A1 according to [ISO 22514-2]

7 Distribution modela A2 and B according to [ISO 22514-2]

7 Distribution models

In this chapter, the density functions of the distributions commonly used in the context of capability studies are presented in order to give an idea of the typical shapes. Details can be found in [Hartung], [Sachs], [Wilrich], [Schulze], for instance.

Note on the usage of language: When speaking of a statistical distribution, this term is usually associated with the graphical representation (the graph) of its probability density function, e. g. the Gaussian bell-shaped curve in 7.1. The probability distribution results from the integration of the probability density function [Booklet 3]. In industrial practice the terms are mostly used synonymously.

Specified characteristics and corresponding distributions

[CDQ 0301] specifies a selection of distributions for certain characteristics, in particular characteristics of form, orientation, location and run-out according to [ISO 1101], which experience has shown to be suitable for describing the associated measurement data. Most of these characteristics can only assume values greater than zero. Consequently, the distribution models assigned to them have a natural lower limit at zero and are also called right-skewed distributions. Examples are the folded normal distribution and the Rayleigh distribution. With them a multiplicity of empirical distributions of these characteristics can be covered.

For one-sided and naturally limited characteristics, see also Appendix K.

7.1 Normal distribution

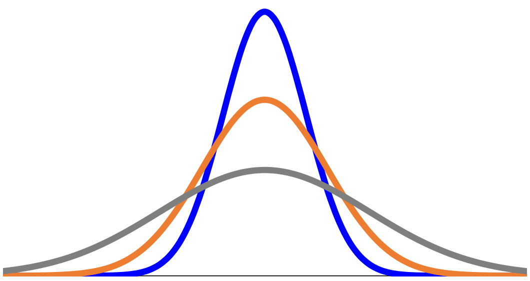


Figure 5: Normal distribution (density functions)

The normal distribution represents statistically the ideal case, which can be handled relatively easily mathematically, but which is often not even approximately achieved in real processes.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

for $-\infty \leq x \leq \infty$

NOTE 1: For capability and performance of processes, numerous standards have been published by international, regional and national standard bodies as well as by the industry. All these standards assume that the considered process is stable and a follows stationary normal distribution. A comprehensive analysis of production processes shows, that processes remain very rarely in such a state when observed over time (in direct reference to [ISO 22514-2], Introduction).

NOTE 2: If there is no normal distribution, it is not permissible to ignore this and, assuming a normal distribution

- *based on the arithmetic mean $\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$ and the empirical standard deviation $s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2}$, make comparisons with specified limits and the corresponding tolerance of a characteristic, or*
- *to calculate limits and a tolerance for a characteristic using $\bar{x}$ and s.*

In particular, it does not make sense to specify symmetrical boundaries $\bar{x} \pm 3 \cdot s$ around the mean value in the case of a skewed distribution.

7.2 Logarithmic normal distribution

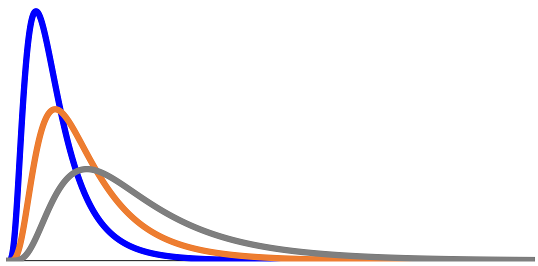


Figure 6: Lognormal distribution (density functions)

The graph of the density function of the logarithmic normal distribution (also called lognormal distribution) shows an asymmetric, zero-limited curve that runs flat on the right.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2} \cdot x} \cdot e^{-\frac{1}{2}\left(\frac{ln(x)-\mu}{\sigma}\right)^2}$$

for $x > 0$

If one takes logarithms of the values of such a distribution, one finds that the results are normally distributed. A continuous random variable X is called logarithmically normally distributed (lognormally distributed), if $Y = ln(X)$ is normally distributed. The left part of the lognormal distribution is strongly stretched by this transformation, the right part is strongly compressed.

7.3 Folded normal distribution

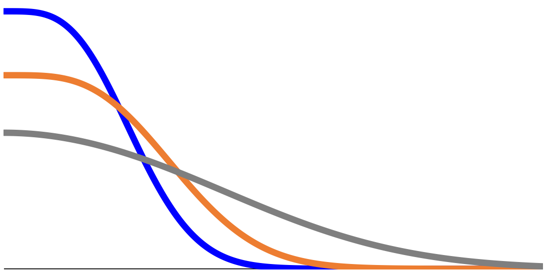


Figure 7: Folded normal distribution (density functions)

A folded normal distribution results, for example, for some zero-limited characteristics of the shape or position, such as straightness, flatness, roundness.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \left[e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} + e^{-\frac{1}{2}\left(\frac{x+\mu}{\sigma}\right)^2}\right]$$

for $0 \leq x < \infty$

Since the target value for such quantities is zero, corresponding measured values will accumulate to the right of the zero point. However, they cannot be less than zero.

For the special case $\mu = 0$ the result is a so-called “half-normal distribution” with the density function

$$f(x) = \frac{2}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \qquad x > 0$$

7.4 Rayleigh distribution

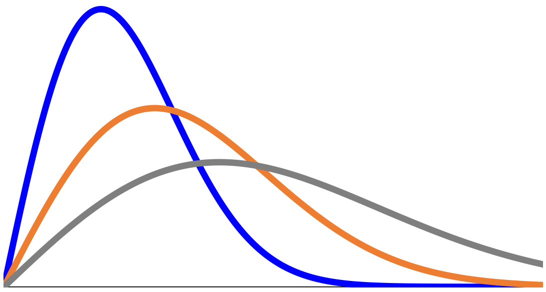


Figure 8: Rayleigh distribution (density functions)

Data that can be described by the Rayleigh distribution result, for example, if a bivariate normally distributed quantity with x and y coordinates is converted into a one-dimensional quantity by calculating the absolute value

$r = \sqrt{x^2 + y^2}$.

$f(x) = \frac{x}{\sigma^2} \cdot e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$ for $x \geq 0$

This applies, for example, to an "unbalance" if polar coordinates with absolute value and direction (angle) are used instead of the x and y coordinates. An example is shown in Figure 25.

7.5 Weibull distribution

The Weibull distribution is a general option for naturally limited characteristics. Although it is often used to evaluate lifetime data, its flexibility in the two- or three-parameter form also makes it suitable for any characteristics that are limited on one side to the left or on one side to the right, provided that no distribution is predefined.

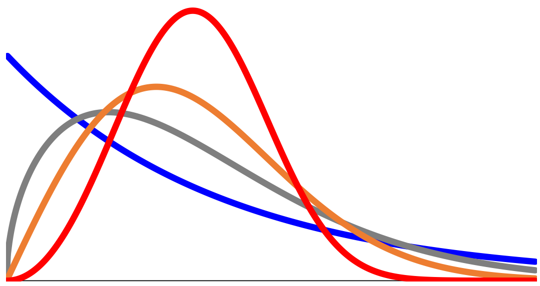


Figure 9: Weibull distribution (density functions)

In the two-parameter form, the Weibull distribution is characterized by the form parameter β and the location parameter (scale parameter) α.

$$f(x) = \frac{\beta}{\alpha} \cdot \frac{x^{\beta-1}}{\alpha} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$$

for $x \geq 0$

The Weibull distribution corresponds to the

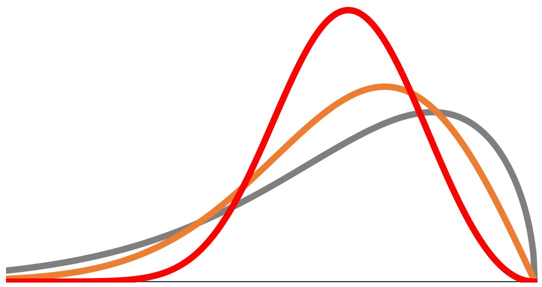
- exponential distribution (blue) if $\beta = 1$ and to the
- Rayleigh distribution (orange) if $\beta = 2$.

For $\beta \approx 3.60235$ the skewness of this distribution is infinitesimally small (red). It resembles the normal distribution but is only defined for $x \geq 0$.

7.6 Distributions for characteristics limited at the upper end

In practice, characteristics that are limited at the upper end are comparatively rare. Examples:

- Pull-off force (of wires or adhesive joints)
- Adhesion (of coatings, lacquers)
- Burst pressure (of membranes)
- Prevail torque or breakloose torque (of screw connections)



In such cases, the measured data show a distribution that is skewed to the left. Such characteristics therefore have a natural limit a, which cannot be exceeded. The right-skewed distributions mentioned in Sections 7.2 and 7.3 can also be used for their description and evaluation, provided that the measured quantity X is transformed into $X^* = \mathrm{a} - \mathrm{X}$. This figure shows an example.

Figure 10: Weibull distribution (density functions) of the quantity $X^* = \mathrm{a} - \mathrm{X}$.

7.7 Mixture distribution

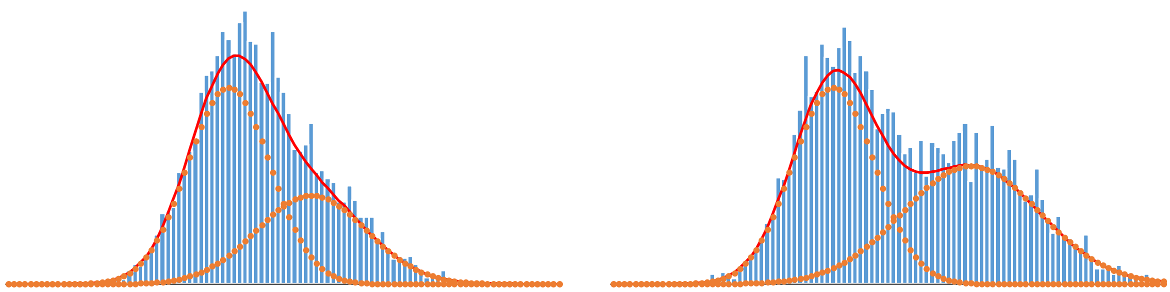
All other distributions considered in this document are unimodal. In contrast, the mixed distribution is a multimodal distribution. It results from the superposition of several normal distributions. In reality, this could correspond, for example, to batchwise production with the same or different quantities of the individual batches and different batch averages.

In order to be able to assign a mixed distribution to an empirical data set, it is necessary to specify the number k of possible peaks (components) for the software used. If this number cannot be given meaningfully on the basis of the technical background of the data, two or three peaks are selected, for example.

The density function of the mixed distribution is composed additively of the density functions $f_i(x, \mu_i, \sigma_i)$ of the individual normal distributions whose weighting is given by the factors a_i.

$f_{mix} = \sum_{i=1}^{k} a_i \cdot f_i(x, \mu_i, \sigma_i)$ with $\sum_{i=1}^{k} a_i = 1$ and $k > 1$

The mean values and variances of the components can be different and the weights a_i correspond to the respective proportion of the total quantity.



The histogram alone would also suggest a single peak skewed distribution.

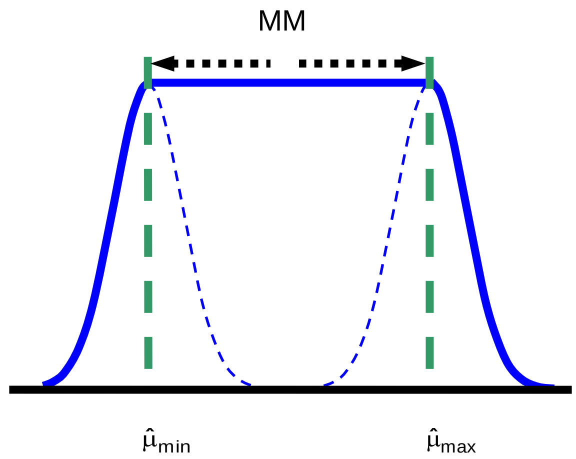
If the mean values differ significantly and the a_i are not too different, the mixed distribution becomes visible in the histogram.

Figure 11: Mixture distributions

Small sample sizes produce histograms with different column heights, which can simulate multiple peaks. The classification model also has an influence on the class assignment and the "appearance" of the histogram, i. e. the relative frequencies above the classes.

The histograms in Figures 15, 17, 18, 24, 32 show certain patterns of the columns which would suggest, for example, the adaptation of a mixed distribution. Whether this could be explained by a technical fact cannot, however, be decided from a purely statistical point of view (see also C.4).

7.8 Extended normal distribution



The graph of the density function of the extended normal distribution shows a strictly symmetric shape.

$$f(x) = f_1(x, \mu_{min}, \sigma) \quad \text{for} \quad x \le \mu_{min}$$
$$f(x) = f_1(\mu_{min}, \mu_{min}, \sigma) \text{ for } \mu_{min} \le x \le \mu_{max}$$
$$f(x) = f_2(x, \mu_{max}, \sigma) \quad \text{for} \quad x \ge \mu_{max}$$

Figure 12: Extended normal distribution (density function)

The left and right flanks correspond to normal distributions with identical standard deviations and mean values separated by the distance MM (moving mean): $\mu_{max} - \mu_{min} = MM$. For the determination of estimated values $\hat{\mu}_{max}$ and $\hat{\mu}_{min}$ see Section 8.4. It is the boundary case of a mixed distribution of infinitely many instantaneous distributions, which is to be expected in reality rather rarely.

The extended normal distribution can arise, when a currently normally distributed process has, in the long term, an additional variation of the mean, e. g. due to tool wear.

This model conception is motivated by a corresponding illustration in [ISO 22514-2], among other things. It can be found there under process model C3, in which, however, both the instantaneous distributions and the resulting distribution have "any shape". As the figure in the standard shows, the shape of the instantaneous distributions remains unchanged over time. With a linear trend, the flanks of the resulting distribution would then have to correspond to those of the instantaneous distribution. In particular, the right flank cannot correspond to that of a normal distribution if the instantaneous distribution is different.

In practice one can assume that the standard deviations σ of the instantaneous distributions do not change by the trend. In qs-STAT® they are therefore assumed to be the same.

7.9 Distributions with offset

For some distributions that are only defined for $x \geq 0$, it is possible to consider a so-called offset parameter. In [Wilrich] it is called a location parameter. It corresponds to a transformation of the type $\mathrm{X}^* = \mathrm{X} - \mathrm{a}$. This clearly means a shift of the density function on the x-axis. This concerns e. g. the lognormal distribution, the absolute distribution, the Rayleigh distribution and the Weibull distribution, whose description formulas then contain three parameters.

7.10 Distributions with convolution

The folded normal distribution occurs when the negative results of a normally distributed population with $\mu \geq 0$ are treated as positive values. Figuratively speaking, the part of the normal distribution to the left of the zero is folded to the right. The originally negative portions in the histogram are thus assigned to the classes to the right of zero.

For the special case $\mu = 0$ and convolution point $a = 0$ the result is a so-called "half-normal distribution". Its density function is greater by a factor of 2 at every point $x \geq 0$ than that of the normal distribution (see 7.3.1).

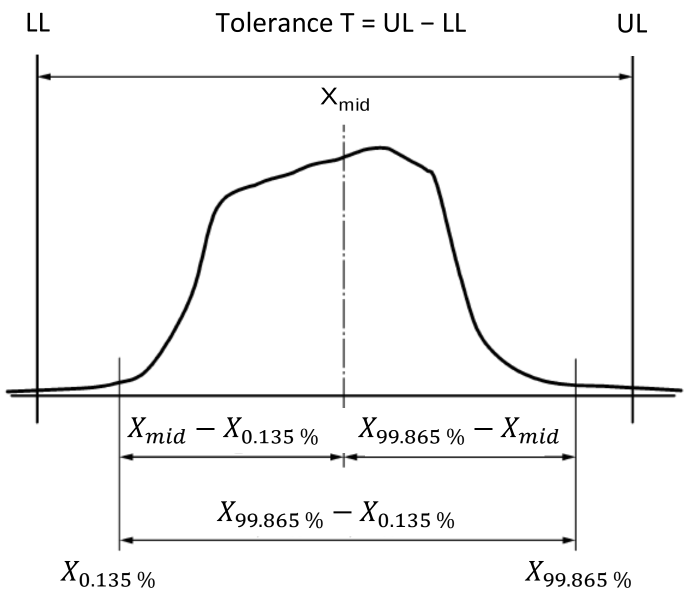
However, such a convolution can also be performed at any position $a > 0$. Details are described for example in [Schulze]. The procedure is used to optimally adapt a distribution function to the existing data set by means of an additional parameter.

> *Note: In the context of statistics, the term "convolution" usually refers to a combination of probability density functions or distribution functions. In the previous section, on the other hand, the term is to be understood rather vividly.*

8 Calculation of capability and performance indices

8.1 Basic principles

As a rule, the distribution model for the measurement values of a product characteristic (the process model) is the basis for determining the statistical parameters *machine capability* (C_m, C_{mk})[8], *production process capability* (C_p, C_{pk}) and *production process performance* (P_p, P_{pk}) as well as their short-term variants (C_{p-ST}, C_{pk-ST}, P_{p-ST}, P_{pk-ST}). The calculation of these parameters is based on the process location X_{mid} *(English **mid**dle)*[9], the variation range of the measured characteristic values x and the specification limits LL *(English "**l**owe**r** limit", German "untere Grenze")* and UL *(English "**u**pper limit")*.



The variation range is limited by generally accepted convention by the quantiles $X_{0,135\,\%}$ and $X_{99,865\,\%}$. Between two quantiles, 99.73 % of all measurements are to be expected, above and below these percentiles respectively 0.135 % of all measurements. For normally distributed measurements, this corresponds to variation range $6 \cdot \sigma$.

Figure 13: Process location and process variation range of any distribution models[10]

For the potential indices C_m, C_p, P_p, C_{p-ST} and P_{p-ST} as well as for the critical indices C_{mk}, C_{pk}, P_{pk}, C_{pk-ST} and P_{pk-ST} same calculation rules apply respectively[11]:

$$\begin{pmatrix} C_m \\ C_p \\ P_p \\ C_{p-ST} \\ P_{p-ST} \end{pmatrix} = \frac{UL - LL}{X_{99,865\,\%} - X_{0,135\,\%}}$$

Equation 8.1

and

$$\begin{pmatrix} C_{mk} \\ C_{pk} \\ P_{pk} \\ C_{pk-ST} \\ P_{pk-ST} \end{pmatrix} = min\left(\frac{X_{mid} - LL}{X_{mid} - X_{0,135\,\%}};\ \frac{UL - X_{mid}}{X_{99,865\,\%} - X_{mid}}\right)$$

Equation 8.2

In the application it depends on the specific task (determining the machine, short- or long-term capability) and the actual process behavior (stable or unstable), which index name is allocated. This assignment does not affect the calculated numerical value.

[8] [ISO 22514-3] uses the variable names P_m and P_{mk} instead of C_m or C_{mk}

[9] [ISO 22514-2] uses the variable name X_{mid} instead of μ commonly used in literature

[10] Diagram directly based on [ISO 22514-2]

[11] *The function min ("minimum") returns the smaller of the two numbers values, which results from the two by semicolon separated calculation rules in the bracket.*

Potential and critical capability and performance indices

The potential indices C_m, C_p, P_p, C_{p-ST} and P_{p-ST} are defined according to the equation (8.1) as the ratio of the widths of the tolerance interval and the random variation of the population which is estimated from the measurements. The potential indices do not contain information about the process location. They are therefore the measurement of the maximum capability or performance of a process, which would be achievable with an ideal centering of the process. They set the technical requirement (tolerance) in relation to an intrinsic property of the process (variation) and are thus a measure of how well the variation of the process and the tolerance interval would be “acceptable” to each other, that is, how well the requirement and property is ideally compatible.

In contrast to this, the critical C_{mk}, C_{pk}, P_{pk}, C_{pk-ST} and P_{pk-ST} also include the process location and are therefore a quantitative measure of how well is the variation of the process *actually* “compatible” with the tolerance interval, that is, “fits” in the tolerance interval.

The greater the difference between potential and associated critical index, the more decentralized is the process location and the greater the likelihood that the measurements are outside the tolerance interval.

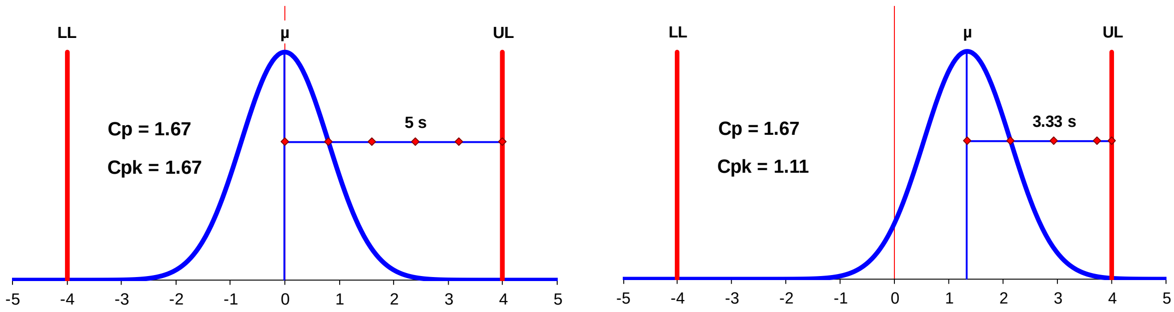


Figure 14: (a) Central process location $C_{pk} = C_p$; (b) decentral process location $C_{pk} < C_p$

Unilaterally restricted characteristics

Only one specification limit (LL or UL) exists for unilaterally limited characteristics. A tolerance which is defined as the difference $T = UL - LL$ does not exist in this case. Thus, no potential index (C_m, C_p, P_p, C_{p-ST} or P_{p-ST}) can be given.

NOTE: It should be noted that this statement also applies if a so-called natural limit exists (for example, LL = 0 for roughness of surfaces, widths of notches and joints). If one would treat this natural limit like a “normal” limit, this would have the result that the process would be aligned as exactly as possible in the middle between the natural and specified limit. Instead, the process in such cases is to be positioned close to the natural limit, so that the distance to the specified limit is as large as possible.*

The rules for calculating the remaining critical index are then reduced to

$$\left.\begin{matrix} C_{mk} \\ C_{pk} \\ P_{pk} \\ C_{pk-ST} \\ P_{pk-ST} \end{matrix}\right\} = \frac{X_{mid} - LL}{X_{mid} - X_{0,135\%}} \qquad (8.3) \qquad \text{and} \qquad \left.\begin{matrix} C_{mk} \\ C_{pk} \\ P_{pk} \\ C_{pk-ST} \\ P_{pk-ST} \end{matrix}\right\} = \frac{UL - X_{mid}}{X_{99,865\%} - X_{mid}} \qquad (8.4)$$

where Eq. (8.3) applies for a unilaterally lower limited characteristic or Eq. (8.4) for a unilaterally upper limited characteristic.

Estimate values

For the calculation of the indices the estimates $\hat{X}_{mid}$, $\hat{X}_{0.135\,\%}$ and $\hat{X}_{99.865\,\%}$ for the location X_{mid} and the quantities $X_{0.135\,\%}$ and $X_{99.865\,\%}$ of the random variation range of the population are required.

[ISO 22514-2] provides 4 methods for estimating the process location (l = 1 … 4, location) and 5 methods for the estimate of the process variation (d = 1 ... 5, dispersion), which can be combined with each other, taking into account the particular characteristics of the data distribution (see Sections 8.2 and 8.3 Appendix E, Table 4). The estimation method used is denoted by Ml,d, whereby for l and d, the type of the chosen method is used (e. g. M2,1).

It is essential for the determination of resilient capability indices to realistically estimate in particular the process variation range on the basis of the the 0.135 % and 99.865 % quantiles $X_{0.135\,\%}$ und $X_{99.865\,\%}$ of the data distribution. This means that primarily the edge regions ("tails") of the distribution are relevant, which means, the best possible estimate of the probabilities of the occurrence of the minimum and maximum values.

The commonly available, from a statistical perspective relatively small data sets (50 values for C_{mk}, 125 values for C_{pk}) usually do not contain sufficient numbers of values in the marginal areas to reliably determine the required quantile. The probabilities for very small and very large values must therefore be estimated by means of a distribution function, which is determined based on frequently occurring values in the mean region around the expected value.

The various distribution functions can be, however, very different in their border areas, so that the estimated variation and thus the corresponding parameters depend strongly on the selected distribution. It is therefore extremely important that the proper distribution is chosen with great care [ISO 22514-1, Chapter 5].

Capability and performance

As already mentioned, solely the stability of the production process determines whether the calculated characteristic values are to be interpreted as capability or performance:

- In case of stable processes we speak of **process capability** and term the characteristics with C_p, C_{pk}, C_{p-ST} and C_{pk-ST}.
- In case of unstable processes we talk of **process performance** and term the characteristics with P_p, P_{pk}, P_{p-ST} and P_{pk-ST}.

The terms have not affect on the numerical value of the statistical value.

Unstable but controlled processes include in particular processes with systematic mean value changes (as a result of continuous trends by tool wear or sudden changes in different batches of material; see also Appendix G).

A meaningful distinction between capability and performance requires a larger data base from a long-term process tracking (as a result of lead time with about 125 values, result of a process capability study, evaluation of several control charts).

8.2 Quantile method $M_{2,1}$ according to ISO 22514-2

The method M2,1 (frequently designated as quantile method) is applicable to all time-dependent distribution models for [ISO 22514-2] and not tied to the at least approximate compliance with certain conditions (such as normal distribution) (see Appendix E, Table 4). It is therefore preferably used for all analyses.

Estimation of the process location:

Type	Calculation	
*l = 2**: $X_{mid} \approx \hat{X}_{50\,\%}$	Adaption of a suitable distribution model to the data set to be analyzed (see Section 6.2.2 and Appendix C) and determination of the 50 %-quantile[12] of this distribution function using appropriate statistic software (for example, qs-STAT®). See Note 1.	
l = 2: $X_{mid} \approx \hat{X}_{50\,\%} \approx \tilde{x}$	$\tilde{x} = x_{\left(\frac{m \cdot n + 1}{2}\right)}$	if $m \cdot n$ is odd
	$\tilde{x} = \frac{1}{2} \cdot \left(x_{\left(\frac{m \cdot n}{2}\right)} + x_{\left(\frac{m \cdot n}{2} + 1\right)} \right)$	if $m \cdot n$ is even

with

x_{ik}	Value no. i in sample no. k *(i = 1 … n; k = 1 … m)*	$x_{(j)}$	Value no. j in the series of all values x_{ik} ordered by increasing size *(j = 1 … m·n)*
n	Number of values per sample *(sample size)*	$\tilde{x}$	Median of all values x_{ik}
m	Number of samples	$\hat{X}_{50\%}$	50 % quantile of the distribution function *(estimate for $X_{50\%}$ of the population)*

*NOTE 1: The calculation type l = 2 * is not part of [ISO 22514-2].*

- *The quantile method estimates the process variation range using the quantiles $X_{0,135\,\%}$ and $X_{99,865\,\%}$ of the distribution function (calculation type d = 1, see below). It would therefore be consistent, to estimate the process location Xmid using the quantile X50 % of this distribution function. This corresponds to the method M13,6 according to [ISO 21747], which is often preset as an option in common statistical software (e. g., qs-STAT ®).*
- *In contrast, [ISO 22514-2] contains no more comparable estimation method. Instead, the median of the measuring values is used for the estimator for X50 % and X50 % on the other hand, as estimator for Xmid. However, this method is not generally applicable. It can unjustifiably lead to significantly more favorable or less favorable results for capability and performance than the method $M1_{3,6}$ according to [ISO 21747]. Deviations in the magnitude of +0.05 and more can e. g. occur when a symmetric distribution function is assigned to a data distribution with statistically insignificant skewness.*
- *The calculation type l = 2 according to [ISO 22514-2] is not recommended for the reasons mentioned. Instead, it is recommended to precede according to the previous method $M1_{3,6}$ according to [ISO 21747]. This type of calculation is denoted in this booklet with l = 2 * (subject to subsequent consideration with a possibly different name in [ISO 22514-2]).*
- *NOTE 2: The use of calculation types l = 1, l = 3 and l = 4 (see Section 8.3) for the process location is according to [ISO 22514-2] also possible under certain conditions (see Appendix E, Table 4; methods M1,1, M3,1, M4,1) are, however, less common in combination with calculation type d = 1 for the process variation range.*

[12] With simpler distribution functions (for example, normal distribution, logarithmic distribution) frequently solvable using inverse distribution functions such as EXCEL worksheet functions NORM.INV, LOGNORM.INV.

Estimation of the process variation:

Type	Calculation	
d = 1:	$\hat{X}_{99.865\,\%} - \hat{X}_{0.135\,\%}$ $\hat{X}_{99.865\,\%} - X_{mid}$ $X_{mid} - \hat{X}_{0.135\,\%}$	Adaptation of a suitable distribution model to the data set to be analyzed (see Section 6.2.2 and Appendix C) and determination of the 0.135 %- und 99.865 %-quantile[12] of this distribution function using appropriate statistical software (for example, qs-STAT®).

with

$\hat{X}_{0.135\,\%}$ 0.135 % quantile of the distribution function
(estimator for $X_{0.135\,\%}$ of the population)

$\hat{X}_{99.865\,\%}$ 99,865 % quantile of the distribution function
(estimator for $X_{99.865\,\%}$ of the population)

NOTE 3: Alternatively the manual, graphical determination by means of probability plot is possible. If the data in the probability plot of the standard normal distribution are not represented by a straight line, the representation in the probability plots of other distributions have to be investigated and the one distribution function that provides the best possible representation of a straight line has to be used.

For very large data sets further, alternative approaches are perhaps possible (see Appendix C.5).

8.3 Further methods according to ISO 22514-2

The applicability of the calculation requirements of this chapter requires in most cases that the data to be evaluated can at least be approximated by a normal distribution with a constant location (see Appendix E, Table 4). In this case, the following simplifications apply:

$X_{mid} \approx \hat{\mu} \qquad X_{99.865\,\%} - X_{0.135\,\%} \approx 6 \cdot \hat{\sigma} \qquad X_{99.865\,\%} - X_{mid} \approx 3 \cdot \hat{\sigma} \qquad X_{mid} - X_{0.135\,\%} \approx 3 \cdot \hat{\sigma}$

Estimate of the process location:

Type	Calculation
$l = 1$:	$\hat{\mu} = \bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$
$l = 3$:	$\hat{\mu} = \bar{\bar{x}} = \frac{1}{m} \cdot \sum_{k=1}^{m} \bar{x}_k$ with $\hat{\mu} = \bar{x}_k = \frac{1}{n} \cdot \sum_{i=1}^{n} x_{ik}$
$l = 4$:	$\hat{\mu} = \bar{\tilde{x}} = \frac{1}{m} \cdot \sum_{k=1}^{m} \tilde{x}_k$

with

x_{ik}	Value no. i in sample no. k *(i = 1 … n; k = 1 … m)*	$\bar{\bar{x}}$	(total) mean of m sample means
n	Number of values per sample *(sample size)*	$\tilde{x}_k$	Median of sample no. k
m	Number of samples	$\bar{\tilde{x}}$	Mean value of the medians of m samples
$\bar{x}_k$	Mean of sample no. k		

NOTE 1: The types l = 1 and l = 3 differ only in the numbering of the data values (l = 1 considers all values as a single sample, l = 3 takes into account the sample structure and, in the case of m = 1 is the same as l = 1). This results in the same total mean value. The distinction is not plausible.

Estimation of the process variation:

Type	Calculation	
d = 2:	$\hat{\sigma} = \sqrt{\overline{s^2}}$ with $\overline{s^2} = \frac{1}{m} \cdot \sum_{k=1}^{m} s_k^2$	$s_k = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_{ik} - \bar{x}_k)^2}$
d = 3:	$\hat{\sigma} = \frac{\bar{s}}{c_4}$ with $\bar{s} = \frac{1}{m} \cdot \sum_{k=1}^{m} s_k$	
d = 4:	$\hat{\sigma} = \frac{\bar{R}}{d_2}$ with $\bar{R} = \frac{1}{m} \cdot \sum_{k=1}^{m} R_k$	$R_k = max(x_{ik}) - min(x_{ik})$
d = 5:	$\hat{\sigma} = s_{total}$ with $s_{total} = \sqrt{\frac{1}{m \cdot n-1} \cdot \sum_{k=1}^{m} \sum_{i=1}^{n} (x_{ik} - \bar{\bar{x}})^2}$	

with

$\overline{s^2}$	Mean value of m sample variances	$\bar{R}$	Mean of ranges of m samples
$\bar{s}$	Mean value of m sample standard deviations	R_k	Range of sample no. k
s_k	Standard deviation of sample no. k	s_{total}	Standard deviation of all measurement values (in all samples)

and

n	2	3	4	5	6	7	8	9	10
c_4	0.798	0.886	0.921	0.940	0.952	0.959	0.965	0.969	0.973
d_2	1.128	1.693	2.059	2.326	2.534	2.704	2.847	2.970	3.078

Table 1: Factors c_4 and d_2 for the sample sizes $n = 2, 3, \ldots, 10$

NOTE 2: The factors c_4 and d_2 depend on the sample size n. c_4 is also designated in older literature as a_n.

*NOTE 3: It is sometimes seen as an advantage that capability and performance characteristics can also be calculated "manually" (as with EXCEL) using the above calculation rules. This, however, **cannot** generally be recommended because in this case, the applicability of the calculations are usually not adequately verified, i. e. the compatibility of the measurement data with a normal distribution. The results would be erroneous and thus meaningless with insufficient normally distributed data.*

NOTE 4: The value of the calculated capability and performance characteristic varies slightly with the type of calculation used to estimate the process location and variation. For the purposes of transparent and accountable results therefore it is not recommended to change a once chosen calculation without a compelling reason.

8.4 Extended normal distribution

The extended normal distribution can be interpreted as a simple special case of a mixed distribution with symmetrical flanks (see Section 7.8, Appendix C.5 and [ISO 22514-2, 6.1.4]).

Alternatives for the determination of MM *(Moving Mean)*

1. Recommended calculation method (standard method qs-STAT® software): Variance analytical determination of the variation of means and hence determination of the variation range MM.

2. Determination based on the variation of the sample means, provided that these are normally distributed: $MM = 6 \cdot \hat{\sigma}_{\bar{x}}$. Thereby designates

   $\hat{\sigma}_{\bar{x}} = s_{\bar{x}} = \sqrt{\frac{1}{m-1} \cdot \sum_{k=1}^{m} (\bar{x}_k - \bar{\bar{x}})^2}$ the standard deviation of the m sample means $\bar{x}_k$.

3. Unless appropriate statistical software is available: Approximate calculation of MM as the difference between lowest and highest process location: Determination of the estimates $\hat{\mu}_{min}$ and $\hat{\mu}_{max}$ for the two extreme positions of the process as averages of the 3 smallest or the 3 largest sample means, meaning, of the 3 samles in extreme situation:

   $\hat{\mu}_{min} = \frac{1}{3} \cdot \left(\bar{x}_{1,min} + \bar{x}_{2,min} + \bar{x}_{3,min}\right)$ $\hat{\mu}_{max} = \frac{1}{3} \cdot \left(\bar{x}_{1,max} + \bar{x}_{2,max} + \bar{x}_{3,max}\right)$

   The parameter MM is clearly the „space", which the systematic averaging change claims, and is therefore estimated as the difference between extreme positions: $MM = \hat{\mu}_{max} - \hat{\mu}_{min}$.

   To estimate the process variation (standard deviation) $\hat{\sigma}$ the calculation methods described in Section 8.3 are available. Thus the performance indices are calculated as follows:

   for bilateral limited characteristics: $P_p = \frac{T - MM}{6 \cdot \hat{\sigma}}$ and $P_{pk} = min\left(\frac{UL - \hat{\mu}_{max}}{3 \cdot \hat{\sigma}}; \frac{\hat{\mu}_{min} - LL}{3 \cdot \hat{\sigma}}\right)$

   for unilateral limited characteristics applies: $P_{pk} = \frac{UL - \hat{\mu}_{max}}{3 \cdot \hat{\sigma}}$ or $P_{pk} = \frac{\hat{\mu}_{min} - LL}{3 \cdot \hat{\sigma}}$

   NOTE 1: Since it is an unstable but controlled process, the determined characteristic values are considered as process performance indices and referred to as Pp and Ppk.

   NOTE 2: For unilateral limited characteristics exists no tolerance T, so that the parameter Pp can not be specified.

   NOTE 3: For unilateral limited, in particular zero-limited characteristics, the measurement results are often not distributed symmetrically around a central value, so that the applicability of symmetrical distribution models such as the extended normal distribution is not given. Here, for example, the procedure described in Appendix C.4 (mixed distribution) could be an appropriate solution.

9 Additional notes to capability indices

In the following are some aspects which anyone who performs or evaluates a machine or process capability should be aware of.

Aim of a capability study is, to derive a statement based on the observed sample results about the process behavior (controlled or not) and a not yet existing population — namely the totality of the parts to be produced in future. This is called a *indirect* or *inductive* conclusion. It is actually assumed that the distribution of the population is already known. Based on *representative* samples are those parameters of this distribution estimated. The term *representative sample* means, that the sample should have all possible properties of the actual or only in the future available population.

In fact is after the determination of the sample values nothing know regarding the temporal stability nor the distribution of values, nor the parameters of the distribution and its timing. All this must be assessed solely on the basis of the few available single values.

9.1 Capability indices and fractions nonconforming

In the literature on process capability is it usually depicted as a direct link between a calculated C_{pk} value and a Fractions Nonconforming, for example, $C_{pk} = 1.33$ corresponds to 32 ppm (unilateral). This relationship is based on the model of the normal distribution and is only in the case of normally distributed data. Differs the real characteristic distribution from the normal distribution (see Chapter 7), then there are as a result usually other Fractions Nonconforming

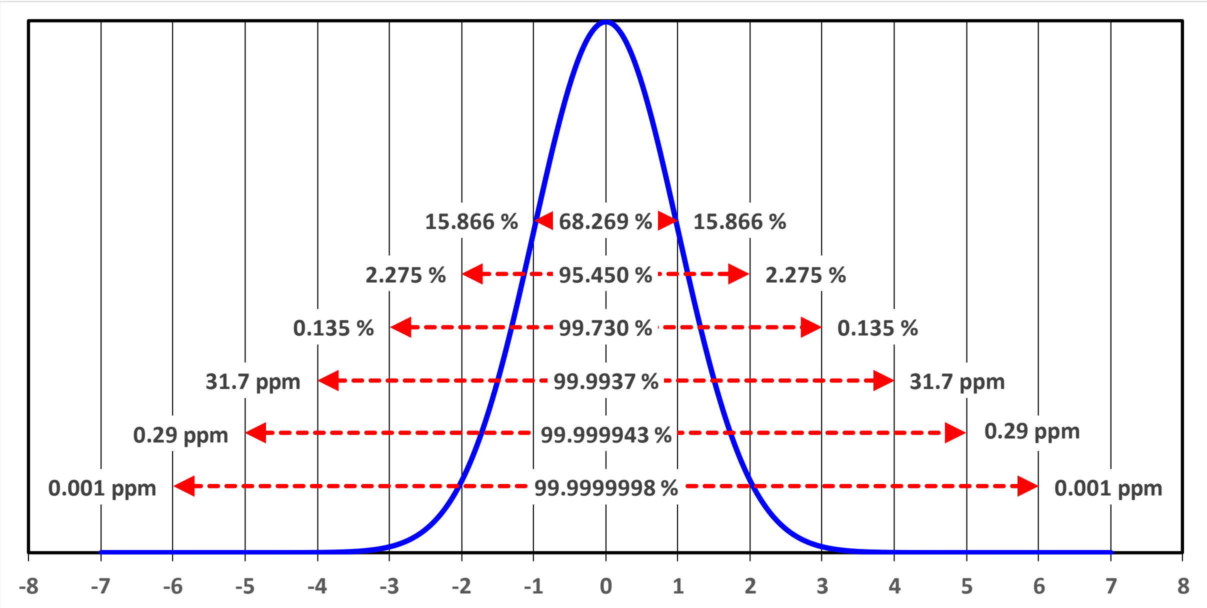


Figure 15: Fractions nonconforming in case of a normal distribution

Within the ranges ±1s, ±2s, ±3s, ±4s, ±5s and ±6s expected proportion of measurement values with associated unilateral Fractions Nonconforming below and above these ranges.

9.2 Influence of the sample size

The sample size has significant impact on the quality of statistical data, which expresses e. g. the size of the confidence intervals estimated distribution parameters such as mean and standard deviation. Statistically, is capability parameter a random variable, which can vary even at an unchanged process from sample to sample.

Appendix C.3 provides additional information on the topic of “confidence interval”, Appendix I to the subject of “Insufficient number of pieces” and specifically Appendix I.1 to the theme “capability and performance limits”

In particular, the allocation of a distribution model (see Appendix D) is more critical, the smaller the sample size. The selection and adaptation of a suitable distribution model is based inter alia on the form of the parameters *skewness* and *kurtosis*. Since these quantities are sensitive to extreme values at a small sample size, a slight change of a few individual values can cause a “switch” in the selection of the distribution model and the corresponding change in the associated capability index.

Appendix C.4 provides additional information on the subject of “distribution selection”.

9.3 Influence of the measurement system

The measuring device and the measuring method used for measuring the components of the sample are of great importance for the later evaluation of the process. Measuring devices with too much uncertainty and unsuitable measuring methods lead to an unnecessary restriction of the tolerance interval for the manufacturing process. A small C_g value or a large %GRR value deteriorates the observed capability and performance characteristics (see Appendix F).

Furthermore is to be noted that a capable measuring device is useless if the parts in the examination are for example dirty, not tempered or are clamped or for example possess great shape deviations.

[Booklet 10] provides the method for determining the “capability of measuring and test processes” with examples and numerous references and additional explanations.

10 Report: Calculated capability and performance indices

To ensure the greatest possible transparency in connection with the detection and dissemination of capability and performance indices at the reporting, the following minimum information should always be provided (see [ISO 22514-2] and Chapter 13):

	Example:
Potential process capability index	$C_p = 1.75$
Critical process capability index	$C_{pk} = 1.47$
Calculation method	$M_{2,1}$
Number of underlying values	200
Optional: • Sampling frequency • Time and duration of data collection • Distribution model (justification) • Measurement system • Technical framework	

11 Revalidation

During the productive use the capability of the manufacturing process must be assured at any time. This is achieved by appropriate time limits for the revalidation of the capability and may be controlled in the time frame between revalidations for example by means of control charts [Booklet 7].

For the interval for regular revalidation of process capability are — as well as on the sampling interval during a single revalidation (see Section 6.1) — no requirements in national and international standards and guidelines.

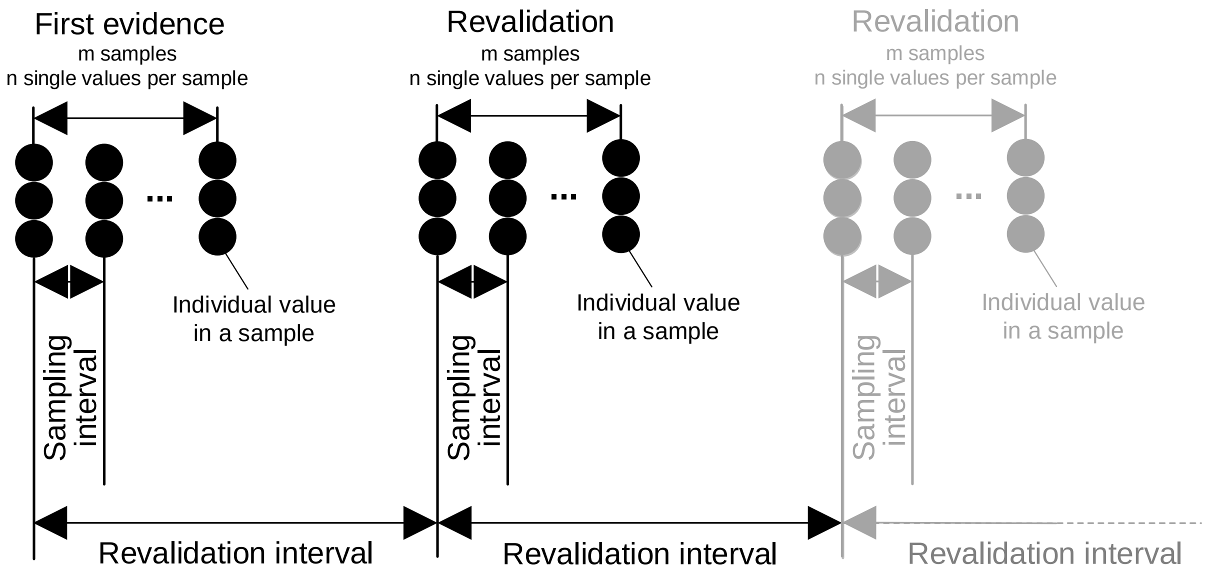


Figure 16: Sampling interval and validation interval (schematically)

Timescales not true to scale: Revalidation intervals are usually much longer than m sampling intervals, which are required for a single revalidation.

The following procedure is a common approach in practice. Three cases have to be distinguished.

Case 1: Controlled processes (SPC processes)

For processes that are monitored by control chart, the most recently acquired data may be accessed at any time to revalidate the process capability. Is so far only the first record for the long-term capability or performance delivered, then there are usually no further informations about the expected future process behavior. In general, it makes sense, to start with the shortest revalidating interval (e. g. evaluation of each or every second control chart) and adapt this interval gradually based on the following observations:

- Over all available control charts no violations of intervention limits occure or not more than the amount that is to be expected by chance.

  NOTE 1: Control limits typically limit the range in which 99.73 % (or 99 %) of all values are to be expected. If a larger number of values is entered, then random, meaning, not process-related violations of the control limits can occur, as 0.27 % (or 1 %) of all values are to be expected outside the control limits.

- From control chart to control chart no systematic changes (for example, “jumps”) can be seen.
- Those control charts, which were each evaluated for validation, provide reproducible capability indices.

NOTE 2: “Reproducible values” means here “values with statistically significant variation”. This “insignificance” can be quantified when necessary by means of statistical tests (see [Booklet 3]). Thus statements are possible such as “The values are with a level of confidence of 95 % indistinguishable”.

If such a behavior is observed consistently over a number of completed control charts, comes an extension (for example, doubling) of the revalidating interval into consideration.

Such adaptation may be repeated several times if the behavior remains unchanged even after extension of the interval. However, it is important to ensure that the interval between two revalidations is not absurd long (e. g. several years). So the production quantity (for example, day, week, month), complexity and criticality of the characteristic in the definition of the interval are suitable examples to be considered.

However, if one of the above criteria is violated, a shortening (for example, halving) of the revalidating interval should be considered.

Case 2: Not controlled processes

If, in a process, which is not monitored by control chart, so far only the first record for the long-term capability or performance is delivered, there are usually no further information about the expected future process behavior. In general, it is useful to carry out at least in the initial phase a control chart and proceed in the same manner as is found in controlled processes to an appropriate revalidation interval.

Case 3: Renewed capability evidence regardless of deadlines

The following criteria are typical examples which make a reanalysis with capability evidence regardless of deadlines inevitable:

- Specification changes of the being manufactured characteristic;
- Increased occurrence of unexpected process results and/or defective parts;
- Intervention in the manufacturing process (for example, after exceeding control limits) lead to process results, which differ significantly from the results, before the intervention was necessary (for example, verifiable on the basis of a control chart);
- Commissioning of new, overhauled or reconditioned production equipment (for example, after maintenance, in which extensive dismantling, rebuilding and/or replacement of essential components were required);
- Technical changes (for example, construction, software), changes of process characteristics (for example, settings) and/or boundary conditions of the manufacturing process (for example, processes, environment);
- Relocation of production equipment.

When in doubt, the analysis must be repeated and the capability has to be proven again.

Since the process changes in these cases, or at least may change, the states before and after the change is often very different. It usually does not make sense to revert to already present measurement data. It is possible that the previous random samples and validation intervals are also no longer adequate.

12 Examples

Example 1: Machine capability

Characteristic:	Disk height in mm
Limits:	$LL = 9.98\ mm$; $UL = 10.00\ mm$; bilateral limited
No. of measurements (parts):	$n = 50$
Calculation method:	Quantile method $M_{2^*,1}$ (with $\hat{X}_{50\ \%}$ of the distrib. function)

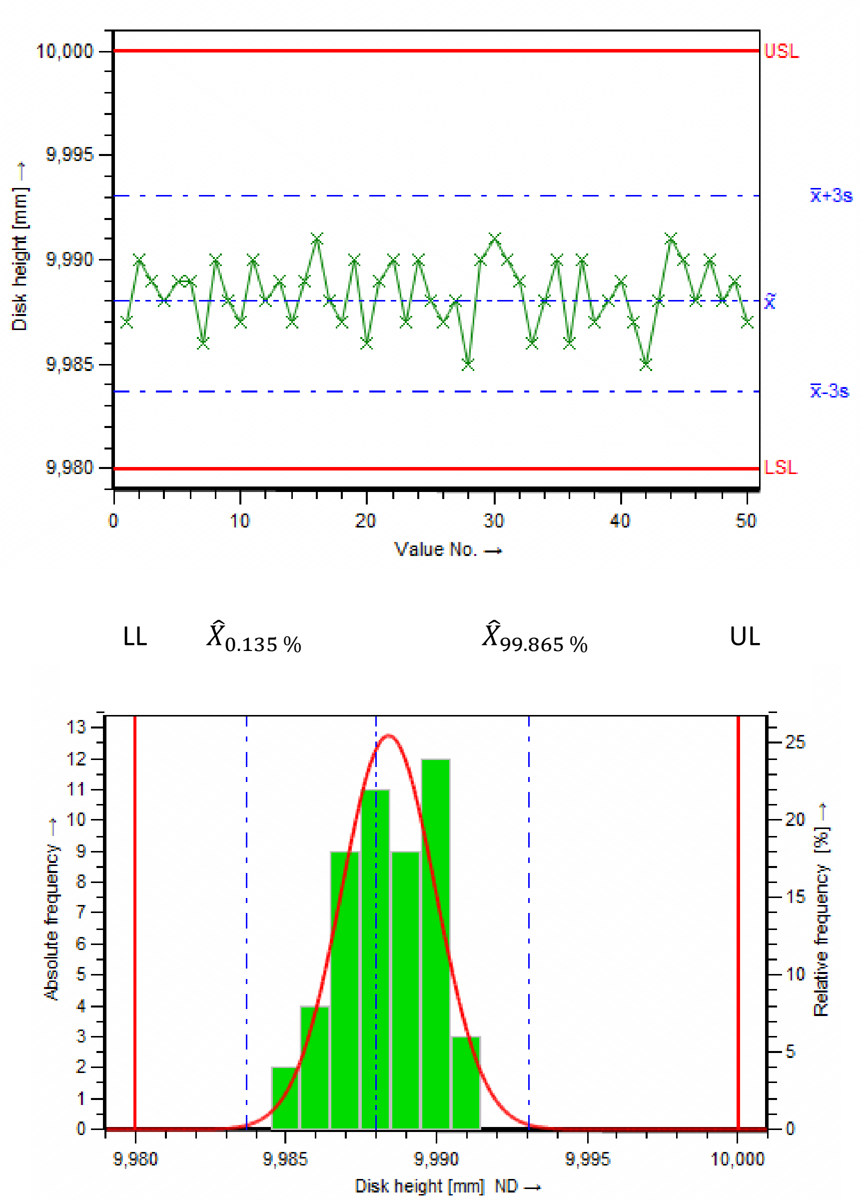


Figure 17: Example "disk height"; Original value chart and histogram with distribution function

Distribution model: Rayleigh distribution

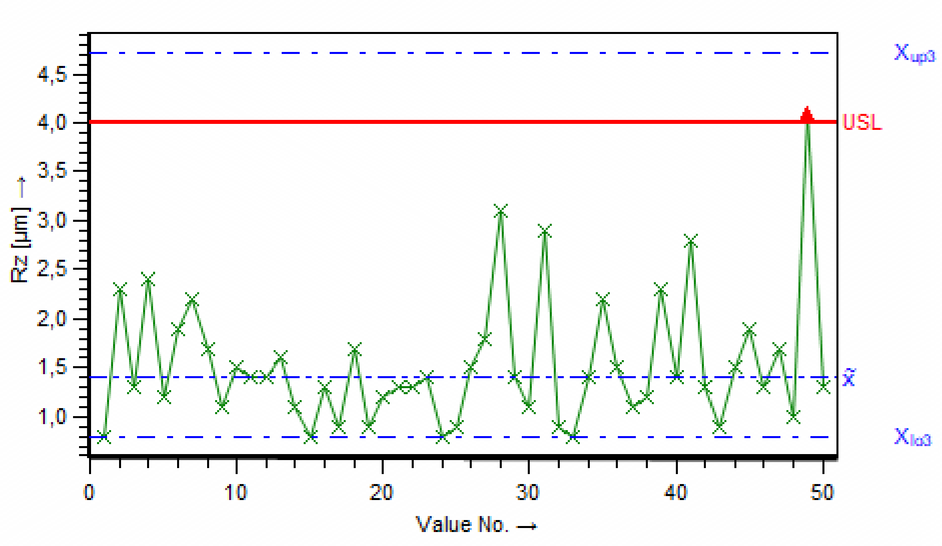
Quantiles: $\hat{X}_{0.135\ \%} = 9.9837\ mm$; $\hat{X}_{50\ \%} = 9.9884\ mm$; $\hat{X}_{99.865\ \%} = 9.9931\ mm$

$$C_m = \frac{UL - LL}{\hat{X}_{99.865\ \%} - \hat{X}_{0.135\ \%}} = 2.13 \qquad C_{mk} = \frac{\hat{X}_{50\ \%} - LL}{\hat{X}_{50\ \%} - \hat{X}_{0.135\ \%}} = 1.79$$

The capability criteria $C_m \geq 1.67$ and $C_{mk} \geq 1.67$ are fulfilled.

Example 2: Machine capability

Characteristic:	Roughness Rz in µm
Limits:	$UL = 4{,}0\ \mu m$; unilaterally limited
No. of measurements (parts):	$n = 50$
Calculation method:	Quantile method $M_{2^*,1}$ (with $\hat{X}_{50\,\%}$ of the distrib. function)



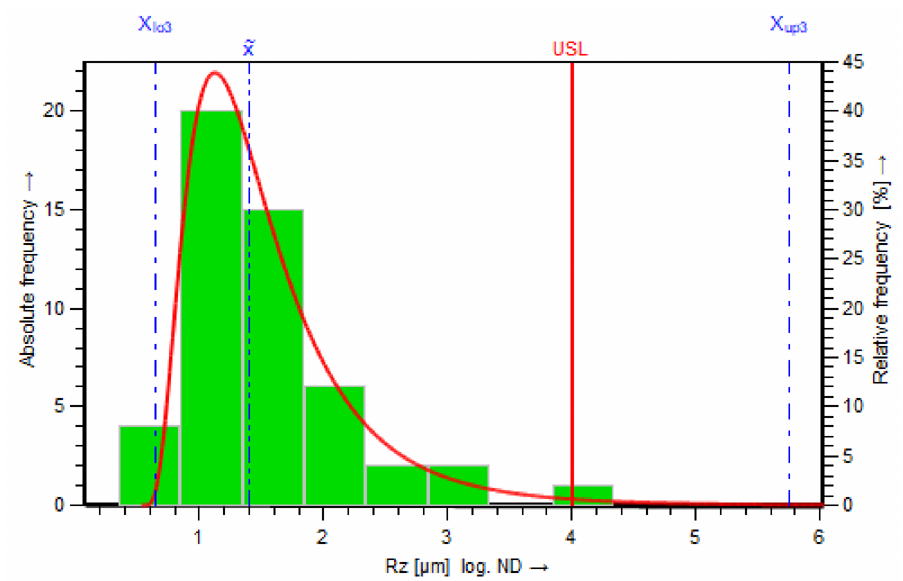


Figure 18: Example „Roughness"; Original value chart and histogram with distribution function

Distribution model: Logarithmic normal distribution

Quantiles: $\hat{X}_{0.135\,\%} = 0.650\ \mu m$; $\hat{X}_{50\,\%} = 1.332\ \mu m$; $\hat{X}_{99.865\,\%} = 5.758\ \mu m$

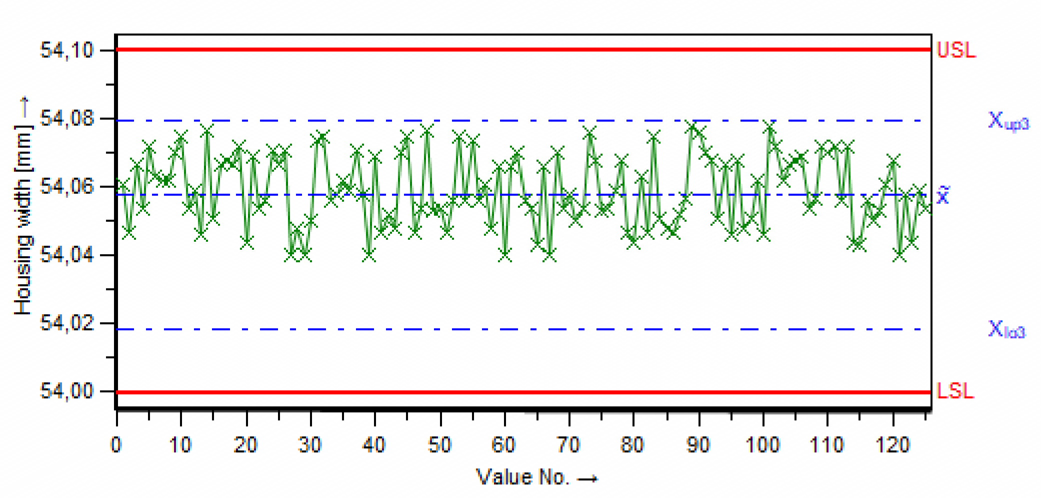
Calculation of C_m does not make sense, since unilaterally limited characteristic

$$C_{mk} = \frac{UL - \hat{X}_{50\,\%}}{\hat{X}_{99.865\,\%} - \hat{X}_{50\,\%}} = 0.60$$

The capability criterion $C_{mk} \geq 1.67$ is **not** fulfilled.

Example 3: Long-term capability of a manufacturing process

Characteristic:	Housing width in mm
Limits:	$LL = 54.0\ mm$; $LL = 54.1\ mm$; bilaterally limited
No. of measurements (parts):	$m \cdot n = 125$
Calculation method:	Quantile method $M_{2^*,1}$ (with $\hat{X}_{50\,\%}$ of the distrib. function)



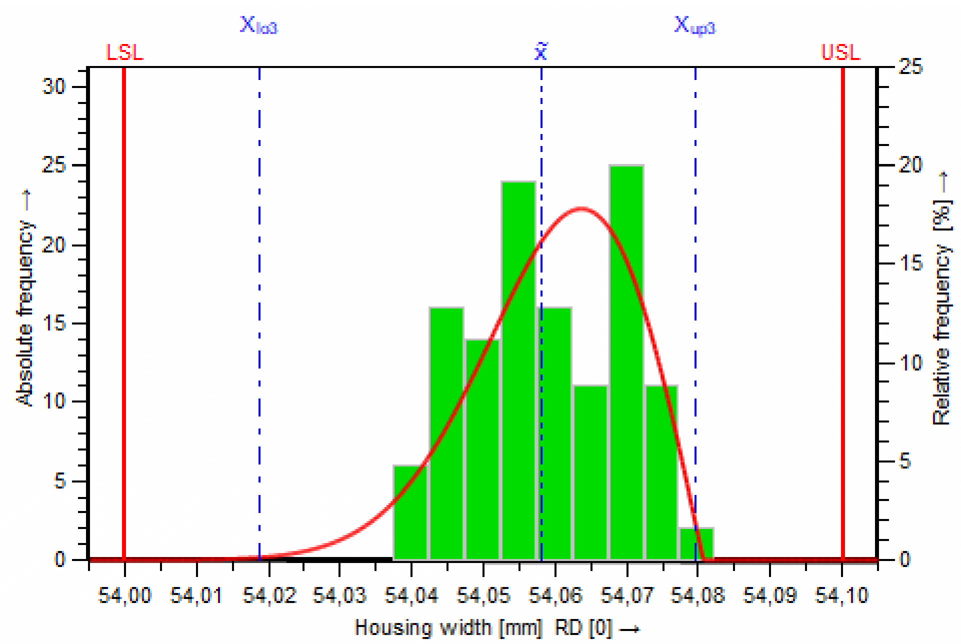


Figure 19: Example „Housing width"; Original value chart and histogram with distribution function

Distribution model: Rayleigh distribution

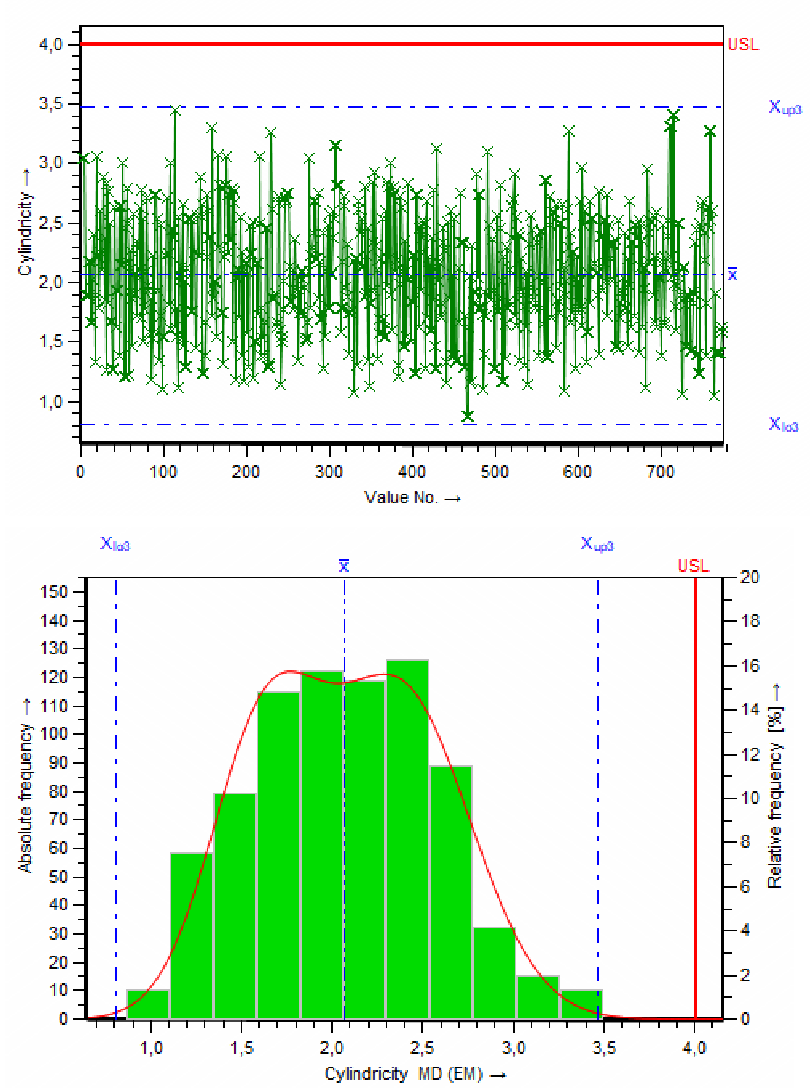
Quantiles: $\hat{X}_{0.135\,\%} = 54.019\ mm$; $\hat{X}_{50\,\%} = 54.280\ mm$; $\hat{X}_{99.865\,\%} = 54.080\ mm$

$$C_p = \frac{UL - LL}{\hat{X}_{99.865\,\%} - \hat{X}_{0.135\,\%}} = 1.64 \qquad C_{pk} = \frac{\hat{X}_{50\,\%} - LL}{\hat{X}_{50\,\%} - \hat{X}_{0.135\,\%}} = 1.45$$

The capability criteria $C_p \geq 1.33$ and $C_{pk} \geq 1.33$ are fulfilled.

Example 4: Long-term capability of a manufacturing process

Characteristic:	Cylindricity in µm
Limit:	$UL = 4.0\ \mu m$; limited at the upper end
No. of measurements (parts):	$m \cdot n = 775$
Calculation method:	Quantile method M2*,1 (with $\hat{X}_{50\,\%}$ of the distrib. function)



Mixture distribution

$$\hat{X}_{0.135\,\%} = 0.81\ \mu m$$
$$\hat{X}_{50\,\%} = 2.06\ \mu m$$
$$\hat{X}_{99.865\,\%} = 3.47\ mm$$

Calculation of P_p does not make sense, since it is a unilaterally limited characteristic.

$$P_{pk} = \frac{UL - \hat{X}_{50\,\%}}{\hat{X}_{99.865\,\%} - \hat{X}_{50\,\%}} = 1.38$$

Figure 20: Example “cylindricity”; Original value chart and histogram with distribution function

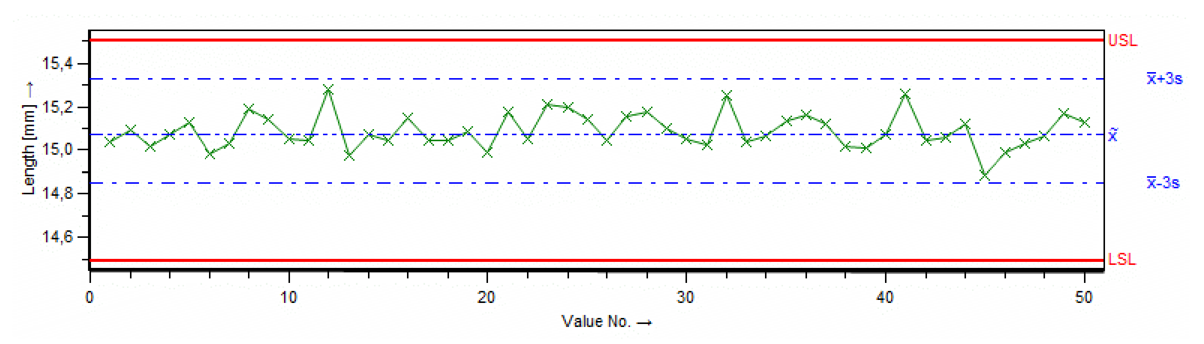
The criterion $P_{pk} \geq 1.33$ for the process performance is fulfilled.

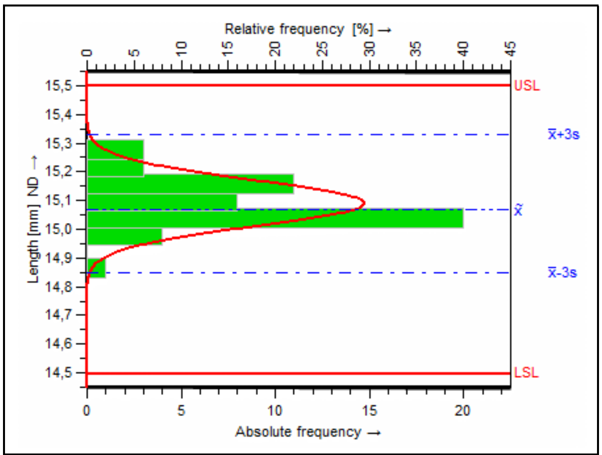
NOTE: Long-term data was analyzed here. Since the process exhibits systematical mean changes, it is not stable in the sense of [ISO 22514-2]. Therefore, the process performance index P_{pk} is indicated.

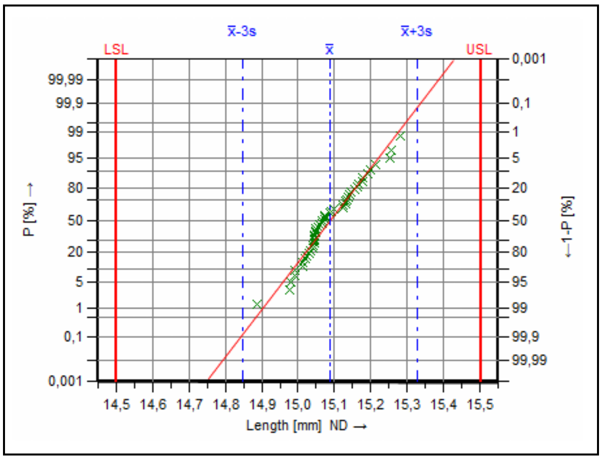
13 Forms

BOSCH	Sample Analysis	Sheet: 1/1

Area	: XyP/MSE9	Operation	: Stamping	Characteristic	: Length
Group/Dptm.	: MOE1.2	Machine	: Punching press	Char. No.	: 2.1
Workshop/sect.	: W123	Machine No.	: 123 456 789	Nomin. Value	: 15.000
Product	: Housing	Test station	: Line 3	Lo. Allowance	: - 0.500
Part	: Cover	Gage	: Calliper	Up. Allowance	: + 0.500
Article number	: X01YZ3A456	Gage No.	: JML012X04-5/10	Tolerance	: 1.000
Change status	: 2019-01-20	Gage Manuf.	: Holex	Unit	: mm
Comment	:				







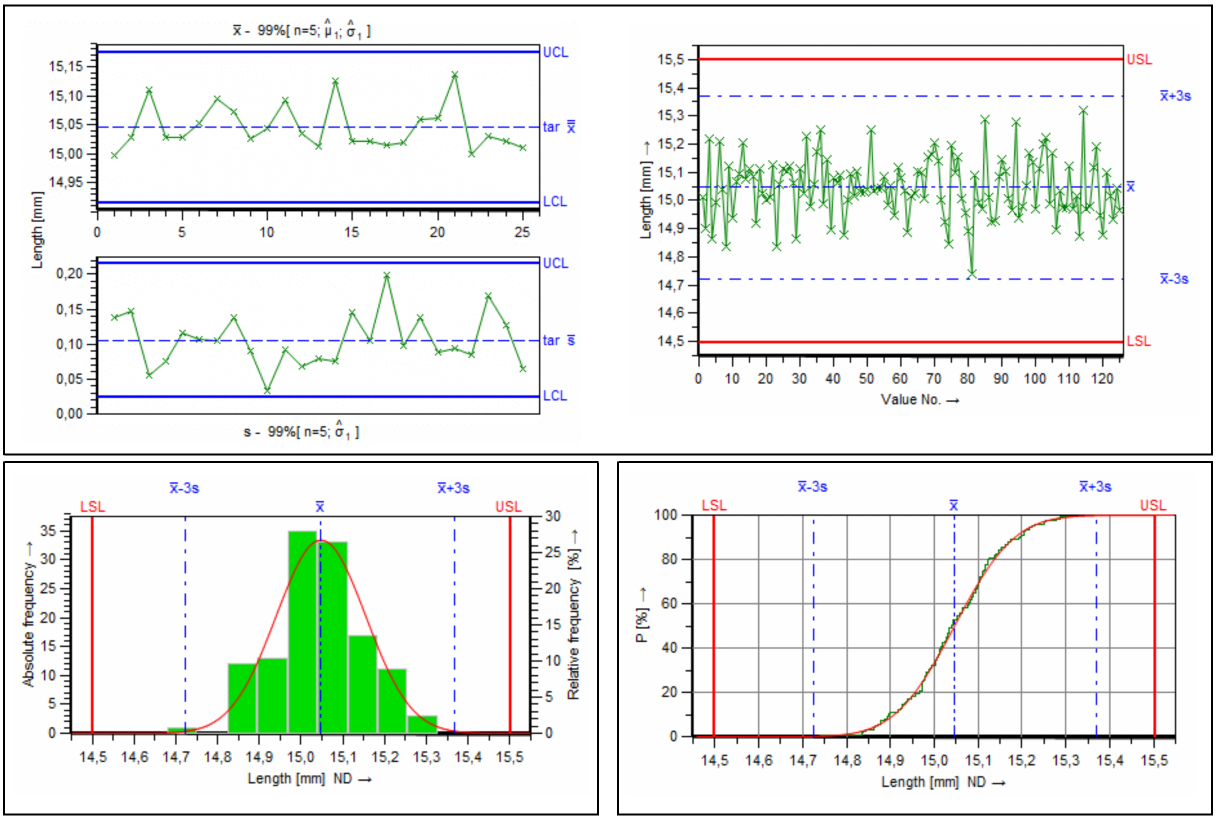
Drawing Values		Collected Values		Statistics	
T_m	15.000	$\bar{x} - T_m$	0,089	$\bar{x}$	15.08910
LSL	14.500	x_{min}	14.887	$\bar{x} - 3\,s$	14.84926
USL	15.500	x_{max}	15.280	$\bar{x} + 3\,s$	15.32894
T	1.000	R	0.393	6 s	0.47969
		$n_{<T>}$	50	$p_{<T>}$	99.99999 %
		$n_{>USL}$	0	$p_{>USL}$	0.00001 %
		$n_{<LSL}$	0	$p_{<LSL}$	0.00000 %
		n_{total}	50	n_{eff}	50

Model distribution		Normal distribution	
Calculation model		$M_{2,1}$ Percentile (0.135%-50%-99.865%)	
Potential Capability Index	C_m	1.67 ≤ **2.08** ≤ 2.50	0 1.67
Critical Capability Index	C_{mk}	1.36 ≤ **1.71** ≤ 2.06	0 1.67
The requirements are met (C_m, C_{mk})			
Evaluation configuration		Bosch 2018	

Figure 21: Contents of the RB standard report on machine capability (qs-STAT®)

NOTE 1: “Random sample Analysis” and “machine capability study” refer to the same investigation.

BOSCH	Process Capability Analysis				Sheet: 1/1
Area	: XyP/MSE9	Operation	: Stamping	Characteristic	: Length
Group/Dptm.	: MOE1.2	Machine	: Punching press	Char. No.	: 2.1
Workshop/sect.	: W123	Machine No.	: 123 456 789	Nomin. Value	: 15.000
Product	: Housing	Test station	: Line 3	Lo. Allowance	: - 0.500
Part	: Cover	Gage	: Caliper	Up. Allowance	: + 0.500
Article number	: X01YZ3A456	Gage No.	: JML012X04-5/10	Tolerance	: 1.000
Change status	: 2019-01-20	Gage Manuf.	: Holex	Unit	: mm
Comment	:				



Drawing Values		Collected Values		Statistics	
T_m	15.000	$\bar{x} - T_m$	0.046	$\bar{x}$	15.04604
LSL	14.500	x_{min}	14.740	$\bar{x} - 3\,s$	14.72383
USL	15.500	x_{max}	15.319	$\bar{x} + 3\,s$	15.36825
T	1.000	R	0.579	6 s	0.64442
		$n_{<T>}$	125	$p_{<T>}$	99.99879 %
		$n_{>OSG}$	0	$p_{>OSG}$	0.00119 %
		$n_{<USG}$	0	$p_{<USG}$	0.00002 %
		n_{total}	125	n_{eff}	125

Model distribution		Normal distribution	
Calculation model		$M_{2,1}$ Percentile (0.135%-50%-99.865%	
Potential Capability Index	C_p	$1.36 \le \mathbf{1.55} \le 1.74$	0 1.33
Critical Capability Index	C_{pk}	$1.22 \le \mathbf{1.41} \le 1.59$	0 1.33
The requirements are met (C_p, C_{pk})			
Evaluation configuration		Bosch 2018	

Figure 22: Contents of the RB standard report on process capability (qs-STAT®)

NOTE 2: See Table8 for special symbols and abbreviations of the Q-DAS software

14 Capability indices for two-dimensional characteristics

The position of a bore hole is an example of a two-dimensional characteristic. The position in a plane is uniquely specified by two coordinates x and y relative to the origin (the point with coordinates (0, 0)). The positional tolerance may be indicated by a circle with radius T/2, corresponding to the center of the target position.

The following statements are intended only to allow an elementary basic understanding, without explaining of the mathematical foundations. Regarding details, please refer to the standard [ISO 22514-6] and corresponding literature.

For the following consideration, it is assumed that the measured positions (x_i, y_i) follow a two-dimensional normal distribution, i. e. both the x and the y component are normally distributed. These positions can be represented in the x-y diagram as points. With the use of appropriate software, a two-dimensional normal distribution will be adjusted to these points and the associated, in this case elliptical random variation calculated, which lies entirely just inside the tolerance circle and includes the proportion $1 - p$ of the population (maximum probability ellipse).

The critical process capability index is then given by $C_{pk} = \frac{u_{1-p}}{3}$.

u_{1-p} designates the $(1 - p)$ quantile of the one-dimensional standard normal distribution. If one shifts the measured points together, so that its center $(\bar{x}, \bar{y})$ coincides with the center of the tolerance circle the elliptical random variation which still lies just completely within the tolerance circle becomes larger and accordingly, the proportion $1 - p'$ contained therein.

The potential capability index is then given by $C_p = \frac{u_{1-p'}}{3}$.

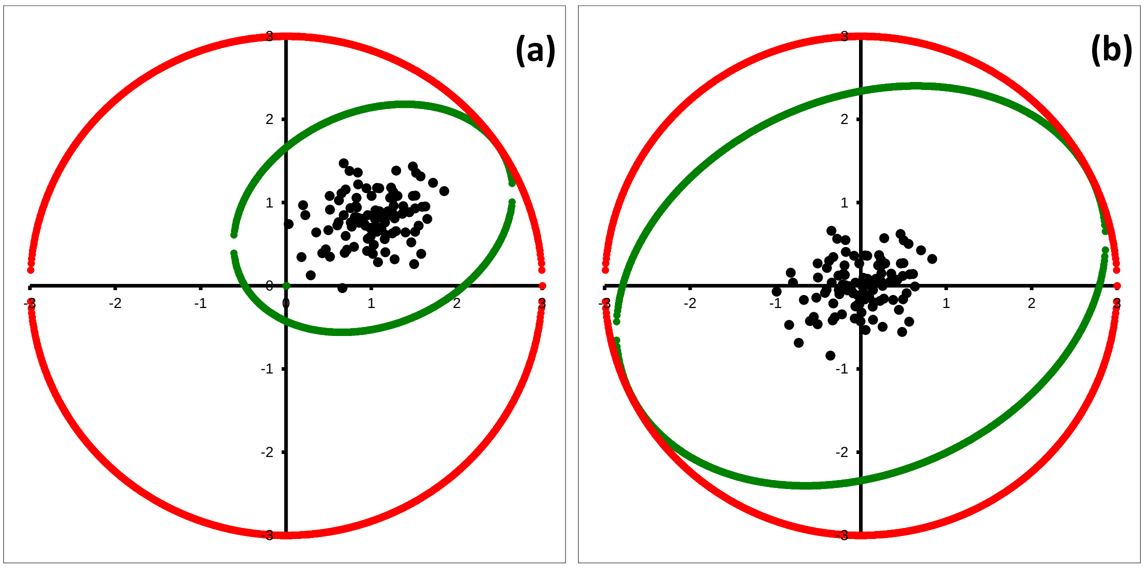


Figure 23: Position measurements, maximum probability ellipse and tolerance circle

(a) The 4s ellipse touches the tolerance circle, $C_{pk} = 1.33$

(b) When centering the point cloud, the 7.6 s ellipse touches the tolerance circle, $C_p = 2.53$.

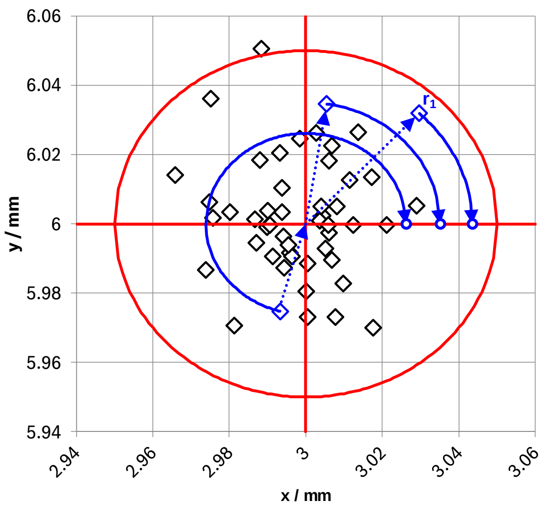
NOTE 1: The outlined method is applicable to any bivariate characteristics and can be generalized using the p-variate normal distribution for characteristics with p components.

A **special case**, which frequently occurs in practice, represents the rotationally symmetrical distribution of position measurements (x_i, y_i) around a reference point (x_0, y_0).

NOTE 2: Reference point may be the specified target position or the center position (mean, center of gravity) of the position measurements, which can be shifted relative to the target position.

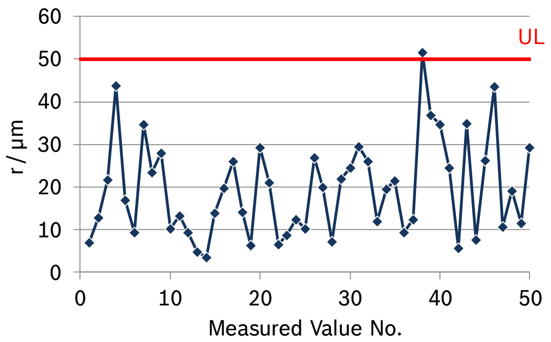
The term "rotational symmetry" means that no significant angular dependence of the radial distances $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ exists, so that the spacings r_i can be regarded as a one-dimensional characteristic and can be evaluated using the Rayleigh distribution (folded normal distribution type 2) e. g. according to method M2,1 (see Section 8.2).

NOTE 3: The Rayleigh distribution is always to be used when the position measurements are described by a two-dimensional, rotationally symmetrical normal distribution. Other distributions are inappropriate in this case and can lead to inappropriate favorable capabilities.



Due to the insignificant pronounced angular dependence in this example, any position measurement (x_i, y_i) while maintaining its distance r_i to the reference point (x_0, y_0) can be rotated randomly around this point of reference, without losing information. These values can therefore be rotated in such way, that after the rotation, all the points on the x-axis lie on the right of the reference point (shown in the graph for 3 measurement values). After rotation, the measurements represent a zero-limited, one-dimensional characteristic with upper specification limit.

Figure 24: Rotationally symmetrical positioned readings (x_i, y_i) with tolerance circle



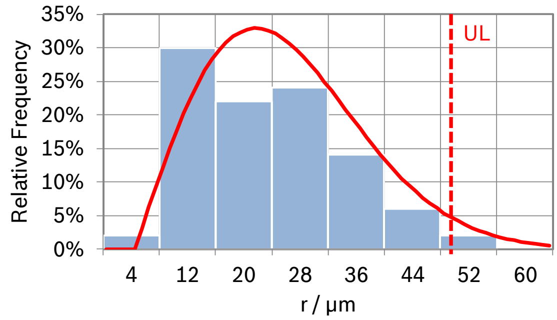


Figure 25: Radial position deviations, histogram and adapted Rayleigh distribution

If the reference point (x_0, y_0) — different from the above example — is not identical to the specified target position, the rotationally symmetrical distribution of the measured position values are no longer concentric in the tolerance circle. In this case, the shortest distance between the reference point (x_0, y_0) and tolerance range for the determination of C_{pk} is authoritative.

Appendix

A Time series analysis

A time series is a data set that contains information for several periods or for several points in time [Friday]. Examples of this are the process capability investigations considered here or data sets obtained within the framework of SPC applications. They usually consist of small samples, e. g. $n = 5$ measured values of a product characteristic, which are determined at short intervals. Such a sample can be understood as a snapshot of the process state and as representative of the instantaneous distribution of a fictitious population.

The temporal development of the "instantaneous distribution" can then be analyzed on the basis of the $j = 1, 2, \ldots, m$ samples, which were recorded at a greater temporal distance from each other. The samples are natural subgroups of the entire dataset of $n \cdot m$ values.

For the tests described in this chapter in the sense of a time series analysis, e. g. with regard to constancy of location or variation as well as possible systematic changes, this grouping in chronological order is an indispensable prerequisite.

A.1 Tests for constancy of process variation

A.1.1 Cochran test

The Cochran test can be used to determine whether the largest of the variances of m samples is significantly different from the variances of the other samples [ISO 5725-2]. Because of the test statistic C, this test is also called Cochran's C-test.

$$C = \frac{s_{max}^2}{\sum_{j=1}^{m} s_j^2}$$

In this expression, s_{max}^2 denotes the largest of the variances s_j^2 of m samples. Prerequisites:

- The samples each have the same size n.
- The individual values x_i within each sample are normally distributed.

Null hypothesis: The variances of the samples are equal.

A.1.2 Analysis of variance (ANOVA) and F-Test

In the one-way analysis of variance, the total variation of all individual values is divided into two parts, a so-called *internal variation* $\overline{s_x^2} = \frac{1}{m} \cdot \sum_{k=1}^{m} s_k^2$ of the individual samples (m groups with $n = 5$ values) and a so-called *external variation* between the individual samples, i. e. the variation of the sample means (cf. Appendix H, Figure 41):

$$s_{\bar{x}}^2 = \frac{1}{m-1} \cdot \sum_{k=1}^{m} (\bar{x}_k - \bar{\bar{x}})^2 \quad \text{with the total mean} \quad \bar{\bar{x}} = \frac{1}{m} \cdot \sum_{k=1}^{m} \bar{x}_k$$

By comparing the test statistic $F = \frac{n \cdot s_{\bar{x}}^2}{\overline{s_x^2}}$ with the critical value of the F-distribution (F-test) for 95 %, it is checked whether the external variation is significantly greater than the internal variation. If this is not the case, the process is considered stable.

NOTE: The significance is determined by the level of confidence, which in this case is usually set at 95 %, i. e. 5 % error probability. The analysis of variance is called "one-way" to distinguish it from a two-way or multiple ANOVA.

A.1.3 Simple alternatives

Appendix B.2 describes in detail a simple stability test for process variation, which can be performed relatively easily "by hand" or e. g. with the help of EXCEL.

A.2 Tests for constancy of process location

A.2.1 Kruskal-Wallis test

The Kruskal-Wallis test (also called H-test) can be used to assess whether the location of the "instantaneous distributions" is constant [Kruskal]. See also Section 6.3. This test is based on the ranks of the individual values ordered by size and does not require a known distribution model. The $N = m \cdot n$ single values $x_{i,j}$ of the m samples of size n are ordered in ascending order. The smallest value has the rank 1, the largest value has the rank N.

R_j is the sum of the ranks of the j-th sample: $R_j = \sum_{i=1}^{n} x_{(i)}$

The test statistic $H = \frac{12}{N \cdot (N+1)} \cdot \frac{1}{n} \cdot \sum_{j=1}^{m} R_j^2 - 3 \cdot (N+1)$ is approximately χ^2 distributed with $m-1$ degrees of freedom. The null hypothesis is rejected if $H > \chi^2_{m-1;\, \alpha}$.

Note 1: [Kruskal] additionally indicates H in a different spelling. It shows that H has the form of a χ^2 distributed random variable. $(N+1)/2$ is the mean, $(N^2-1)/12$ is the variance of ranks 1, 2, ..., N.

Note 2: In the above description it is assumed that the data set does not contain identical values $x_{i,j}$, i. e. the ranks are all different. Otherwise, a correction of the test variable may be necessary. Details are described e. g. in [Wilrich], [Sachs] and [Kruskal].

Prerequisites:

- Grouped data are available, whereby the individual groups are regarded as independent samples from a "instantaneous distribution".
- All samples are subject to the same distribution form. The distribution function is continuous.

Null hypothesis: The m samples originate from the same population.

A.2.2 Simple alternatives

Appendices B.1 and B.3 describe in detail some variants of stability tests for the process location, which can be performed relatively easily "by hand" or, for example, with the aid of EXCEL.

A.3 Test for trend

A simple trend test according to [Neumann] uses the sum of squared differences of successive values $\Delta^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - x_{i+1})^2$ compared to the variance $s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$ of the data set. If the successive values are independent of each other $\frac{\Delta^2}{s^2} \cong 2$. A present trend leads to $\Delta^2 < 2 \cdot s^2$. The reason for this is that s^2 becomes larger by the trend, while it hardly influences Δ^2.

This method of successive differences works for the detection of sawtooth-like time series (trends) as well as jerky, meander-shaped changes with temporal horizontal courses (batch jumps).

The order of the $(x_i - \bar{x})^2$ does not play a role in the expression for variance. The test assumes that the x_i come from a normally distributed population. For details see [Neumann], [Sachs] and [Schulze].

A.4 Test for randomness

In the Swed-Eisenhart [Swed] version, the test generally refers to a characteristic that can only assume two different results, e. g. the results heads "H" and tails "T" for coin tossing. Such a characteristic is called dichotomous. Since the probability for the results "H" and "T" is ½, it is expected that a sequence of several identical results, a so-called "run", will occur rather seldom in successive tosses and the longer the run, the rarer it will be. The probability for the run "HHHHHHH" is $\left(\frac{1}{2}\right)^7 \approx 0.0078$. So it only occurs in less than 1 % of the cases.

For a continuous characteristic, this principle can serve as a randomness test. Regardless of the distribution model, there are the same number of values below and above the median. In a consecutive series of measured values, longer runs of values that all have only positive or only negative deviations from the median should therefore occur very rarely. The run test according to Wald-Wolfowitz uses the number of runs in a value sequence to check whether the order of the values is random or not. This allows trends or periodicities to be identified in a chronological sequence of values.

B Simple statistical stability tests

Criterion for sufficient stability on the confidence level $(1-\alpha)\cdot 100\ \%$ are values within the $(1-\alpha)\cdot 100\ \%$ random variation range of the population, from which these values originate. The boundaries of this area are called control limits in quality control charts.

The standard deviation σ_x of the individual values in the population is going to be

$$\sigma_x \approx \hat{\sigma}_x = \frac{\bar{s}}{c_4} \tag{B.1}$$

estimated with

$\bar{s}$	Mean value of the standard deviations of all samples
n	Number of individual values per sample (sample size)
m	Number of random samples
c_4	Factor according to Table 1

The standard deviation $\sigma_{\bar{x}}$ of the sample mean values $\bar{x}$ is estimated from the individual values according to

$$\sigma_{\bar{x}} \approx \hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}_x}{\sqrt{n}} = \frac{\bar{s}}{c_4 \cdot \sqrt{n}} \tag{B.2}$$

whereby $\hat{\sigma}_x$ was replaced according to Eq. (B.1).

B.1 Location of the individual samples

Due to the central limit theorem of statistics the means $\bar{x}_j$ of samples, each consisting of $n \geq 5$ individual values, may be considered as normally distributed. Accordingly, the quantities $\frac{\bar{x}_j - \bar{\bar{x}}}{\sigma_{\bar{x}}}$ are to be expected with probability $(1-\alpha)\cdot 100\%$ in the interval $[u_{\alpha/2}; u_{1-\alpha/2}]$ of the standard normal distribution.

Related to the original scale of the measurement system this corresponds to a normal distribution with mean $\bar{\bar{x}}$ (mean of the sample means $\bar{x}$) and standard deviation $\sigma_{\bar{x}}$

$$\bar{\bar{x}} - u_{\alpha/2} \cdot \sigma_{\bar{x}} \leq \bar{x}_j \leq \bar{\bar{x}} + u_{1-\alpha/2} \cdot \sigma_{\bar{x}} \tag{B.3}$$

$\sigma_{\bar{x}}$ replaced by Eq. (B.2) results in

$$\bar{\bar{x}} - \frac{u_{\alpha/2}}{c_4 \cdot \sqrt{n}} \cdot \bar{s} \leq \bar{x}_j \leq \bar{\bar{x}} + \frac{u_{1-\alpha/2}}{c_4 \cdot \sqrt{n}} \cdot \bar{s} \tag{B.4}$$

The lower and upper limit for $\bar{x}$ correspond to those control limits on an $\bar{x}$ chart.

Results for sample size $n = 5$ and the confidence levels 99 % and 99.73 %:

$\alpha = 0.01$: $\quad \bar{\bar{x}} - \dfrac{2.58}{0.94 \cdot \sqrt{5}} \cdot \bar{s} \leq \bar{x}_j \leq \bar{\bar{x}} + \dfrac{2.58}{0.94 \cdot \sqrt{5}} \cdot \bar{s}$ or $\bar{\bar{x}} - 1.23 \cdot \bar{s} \leq \bar{x}_j \leq \bar{\bar{x}} + 1.23 \cdot \bar{s}$

$\alpha = 0.0027$: $\quad \bar{\bar{x}} - \dfrac{3.00}{0.94 \cdot \sqrt{5}} \cdot \bar{s} \leq \bar{x}_j \leq \bar{\bar{x}} + \dfrac{3.00}{0.94 \cdot \sqrt{5}} \cdot \bar{s}$ or $\bar{\bar{x}} - 1.43 \cdot \bar{s} \leq \bar{x}_j \leq \bar{\bar{x}} + 1.43 \cdot \bar{s}$

The **process location** is considered as stable with a level of confidence $(1-\alpha)\cdot 100\ \%$, when the mean values $\bar{x}_j$ of the individual samples is between the associated limits.

B.2 Variation range of the individual samples

The distribution of the standard deviations s_i of the sample mean values is equally described by the χ^2 -distribution. Accordingly, the standard deviations with probability $(1-\alpha)\cdot 100\,\%$ can be expected in the interval, which is limited by the corresponding quantile of the χ^2 distribution:

$$\sqrt{\frac{\chi^2_{n-1;\alpha/2}}{n-1}}\cdot\sigma_x \le s_i \le \sqrt{\frac{\chi^2_{n-1;1-\alpha/2}}{n-1}}\cdot\sigma_x \tag{B.5}$$

The estimate $\hat{\sigma}_x$ according to Eq. (B.1) instead of σ_x used, gives the upper limit

$$s_i \le \sqrt{\frac{\chi^2_{n-1;1-\alpha/2}}{n-1}}\cdot\frac{\bar{s}}{c_4} \tag{B.6}$$

This upper limit corresponds to the upper control limit of an s-chart.

Results for sample size $n=5$ and the confidence levels 99 % and 99.73 % [13]	$\alpha = 0.01$:	$s_i \le \sqrt{\frac{14.86}{5-1}}\cdot\frac{\bar{s}}{0.94}$	or	$s_i \le 2.05\cdot\bar{s}$
	$\alpha = 0.0027$:	$s_i \le \sqrt{\frac{17.80}{5-1}}\cdot\frac{\bar{s}}{0.94}$	or	$s_i \le 2.24\cdot\bar{s}$

The **process variation** is considered as stable with a level of confidence $(1-\alpha)\cdot 100\,\%$, when the standard deviations s_i of the individual samples is below the upper limit.

B.3 Standard deviation of the sample means

The distribution of the standard deviations $s_{\bar{x}}$ of the sample means is equally described by the χ^2 distribution. However, now the m individual samples have to be consider as a single sample of the size m:

$$\sqrt{\frac{\chi^2_{m-1;\alpha/2}}{m-1}}\cdot\sigma_{\bar{x}} \le s_{\bar{x}} \le \sqrt{\frac{\chi^2_{m-1;1-\alpha/2}}{m-1}}\cdot\sigma_{\bar{x}} \tag{B.7}$$

Inserting the estimate $\hat{\sigma}_{\bar{x}}$ according to Eq. (B.2) instead of $\sigma_{\bar{x}}$ gives the upper limit

$$s_{\bar{x}} \le \sqrt{\frac{\chi^2_{m-1;1-\alpha/2}}{m-1}}\cdot\frac{\bar{s}}{c_4\cdot\sqrt{n}} \tag{B.8}$$

Results for $m=25$ samples with sample size n and the confidence levels 99 % and 99.73 % [13]:	$\alpha = 0.01$:	$s_{\bar{x}} \le \sqrt{\frac{45.558}{25-1}}\cdot\frac{\bar{s}}{c_4\cdot\sqrt{n}}$	or	$s_{\bar{x}} \le \frac{1.38\cdot\bar{s}}{c_4\cdot\sqrt{n}}$
	$\alpha = 0.0027$:	$s_{\bar{x}} \le \sqrt{\frac{50.163}{25-1}}\cdot\frac{\bar{s}}{c_4\cdot\sqrt{n}}$	or	$s_{\bar{x}} \le \frac{1.45\cdot\bar{s}}{c_4\cdot\sqrt{n}}$

The **process location** is considered as stable with a level of confidence $(1-\alpha)\cdot 100\,\%$, when the standard deviation of sample means $\bar{x}_j$ is below the upper limit.

13 When using the EXCEL function CHIINV, please note that $\alpha/2$ is to be used instead of $1-\alpha/2$ as probability.

C Measurement results and distribution models

NOTE: For the considerations of this chapter only the total count data is relevant, not their allocation to individual samples. Therefore all measurements are treated as one sample of size n.

C.1 Assignment of distribution models

Acquired measurements can i. a. be presented in the form of so-called original value charts (see Figure 26a). However, to statistically evaluate these measurement results, it is necessary to find and adapt an appropriate distribution model (see Figure 26b). Numerous possibilities exist for this: This chapter explains the basics of so-called quantile-quantile plots, which in particular, are frequently used in approximate normally distributed data, but are not limited thereto.

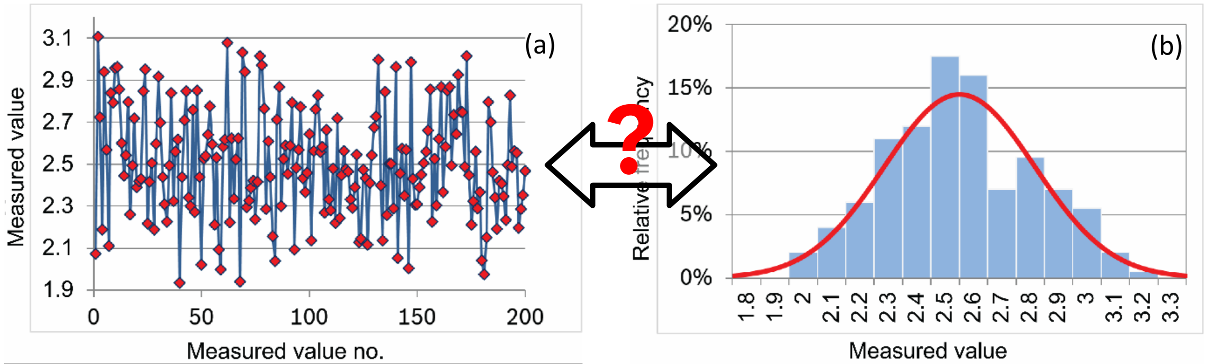


Figure 26: (a) Original value diagram; (b) associated histogram with adjusted distribution density

Basic idea and concept

- These values represent the p-quantile of an unknown distribution.
- The appropriate distribution is determined by comparing the p-quantile of the unknown distribution with the p-quantile of known distributions (quantile-quantile plot).

Meaningful distributions are selected primarily based on technical constraints (for example, limitations of characteristic values due to technical reasons) and possibly existing specifications. Aproach:

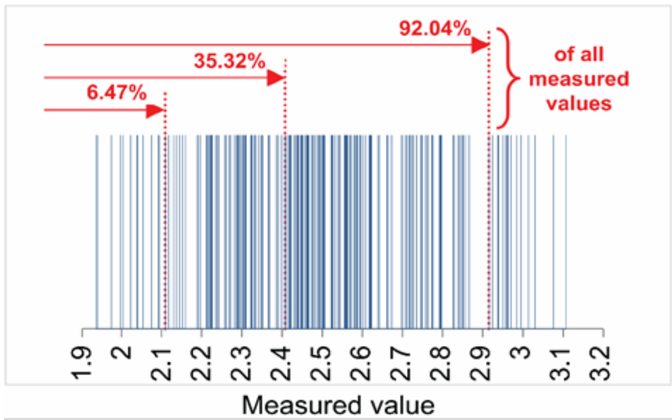


Figure 27: Measurements x in ascending order

- Order the measurements x_i by ascending size.
- Determine for each value x_i the number $\nu(x_i)$ of measurements in the range $\leq x_i$. This is the number of the readings that are "left" of the respective measurement x_i.
- Determine the relative proportions $p(x_i)$ by dividing the figures obtained by the total of all n measurements: $p(x_i) = \frac{\nu(x_i)}{n}$.

- The proportions $p(x_i)$ above the measurements x_i are plotted as a so-called cumulative curve, corresponding to a rough estimate of the distribution curve.

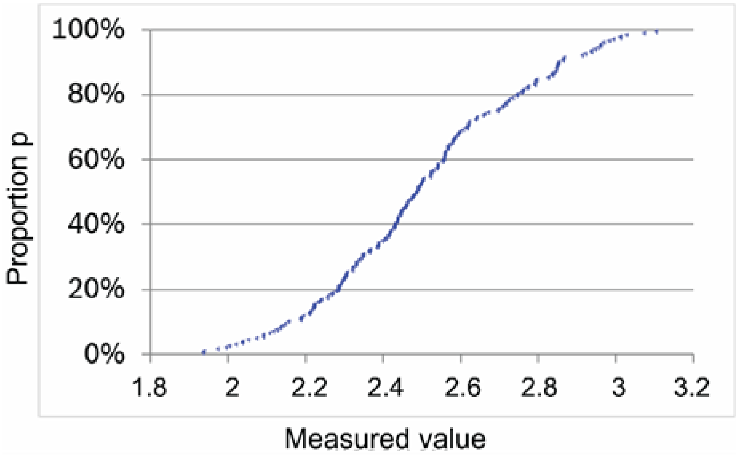


Figure 28: Proportions p(x) applied over the measurements x

Example of normal distribution: .

- The scale of the y-axis of the sum curve $p(x_i)$ is rescaled using the known distribution curve $p(u_p)$ of the standard normal distribution (see Figure 29 a) in the scale of the quantile u_p of the standard normal distribution (quantile-quantile plot, see Figure 29 b).

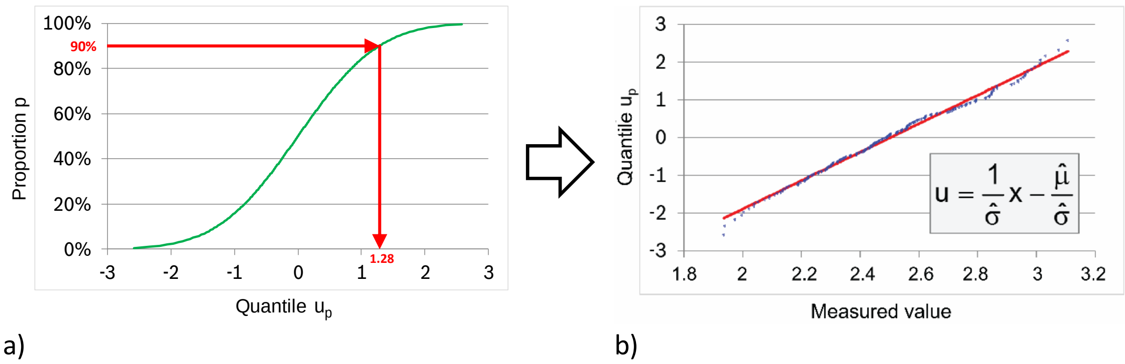


Figure 29: Rescaling of proportions p in quantiles u_p for the example of a normal distribution

- In the ideal case of normally distributed measuring values x_i, all points (x_i, u_p) would lie exactly on a straight line.
- In real measurements, this is not the case in general. Instead, a best fit line $u(x)$ is determined using a linear regression, which provides the parameters $\hat{\mu}$ and $\hat{\sigma}$.

When compared to other distribution models, the procedure is analogous.

C.2 Selection of distribution models

In practice, often multiple, different distribution models can be well adapted to the measurement data, without significant differences in the goodness of fit e. g. visually recognizable. Therefore, a quantitative selection criterion is needed. Typical examples:

- Distribution with the largest correlation coefficient r:

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x}) \cdot (y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \cdot \sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

  The x_i are the measurements and y_i the respective quantiles of the distribution model against which it is compared (i. e., the quantile u_P in Figure 29 b which is assigned to the measurement x_i). If specification limits are set: Correlation coefficient $r_{25\,\%}$ calculated from the 25% of all measurements which are closest to the critical limit (option e. g. available in the Q-DAS software).

- χ^2-test

The following example shows the correlation coefficients which arise with the same record (see Figure 29 a) for different distribution models:

Distribution model	**r** ***(r100 %)***	**r25 %**
Normal distribution	**0.99745**	**0.94724**
Folded normal distrib.	0.48309	
Rayleigh distribution	0.64373	
Weibull distribution	0.98936	0.96603

Table 2: Example of correlation coefficients of various distribution models

In this example, the normal distribution provides the (mathematically) best adaptation.

NOTE: It is expressly noted that this is a purely mathematically justified result. Whether the determined distribution is actually compatible with the technical constraints can be assessed only by the user.

C.3 Parameter estimation and confidence interval

The basic task in determining distribution models is to deduce the (usually unknown) characteristics of the population from which the samples were taken on the basis of representative samples. The term "representative" means that the characteristics of the basic population must be included in the sample as completely as possible so that a meaningful conclusion can be drawn from the sample to the basic population.

Unfortunately, there is no criterion which allows to determine whether a sample is sufficiently representative of the population or not. In addition, the characteristics of the population are often conceptually not fully known, since they do not yet fully exist.

> *NOTE: That's the rule in production processes. Based on samples from the already produced proportion of the population (total of all previously produced product components) the characteristics of the future production proportion of the population will be concluded.*

To come as close as possible to the representative samples, the basic requirement of statistics should be with regards to randomness be adequately met. This means, that sampling elements of the population must be taken at random and may not be specifically selected.

Only then is to expect that the sample elements with certain probability will be within a range of values corresponding to the (generally unknown) variation of the population and is referred to as **random variation range**.

The empirical characteristics of the sample are determined based on these sample elements. These include at least mean value $\bar{x}$ and standard deviation s as variables of location or variation. For non-normally distributed data more parameters are required (for example, skewness, kurtosis).

The determined empirical sample characteristics (e. g. $\bar{x}$ and s) are used as an estimate ($\hat{\mu}$ or $\hat{\sigma}$) for the corresponding parameters of the population (μ or σ), meaning, in the simplest case

$\mu \approx \hat{\mu} \approx \bar{x}$ and $\sigma \approx \hat{\sigma} \approx s$.

This procedure is referred to as **indirect inference** (see Figure 30).

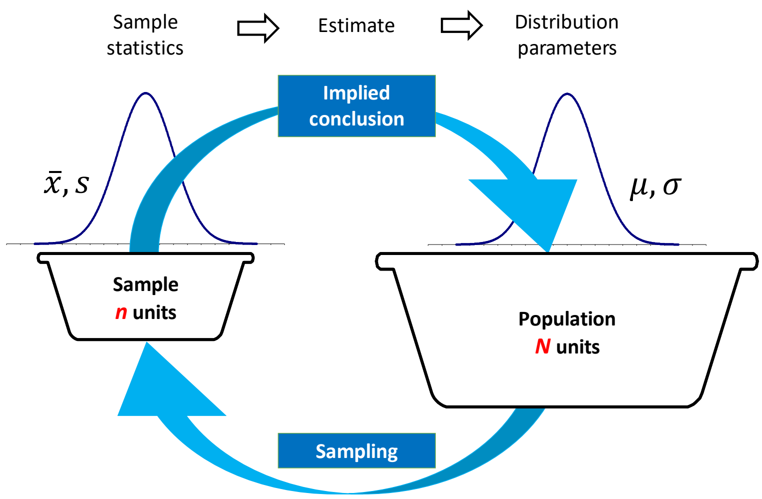


Figure 30: Indirect inference — Statistics, estimators, parameters

NOTE: Usually ***several*** *samples are drawn and estimators for μ and σ are determined according to specific calculation instructions, such as e. g.*

$\mu \approx \hat{\mu} \approx \bar{\bar{x}}$ *and* $\sigma \approx \hat{\sigma} = \sqrt{\overline{s^2}}$

(method $M_{3,1}$);

see Chapters 8.2 and 8.3.

Each estimate is connected to an uncertainty. The "reliability" of the determined parameters of the population is evaluated by the so-called **confidence area**:

> Area in which, on the basis of determined **sample characteristics** (mean, variance) and the distribution model (for example, normal distribution), the true value of a **particular parameter of the population** is to be expected with a given probability (confidence level, degree of confidence).

The meaning is illustrated in the following example:

- There are drawn **hypothetically** 100 samples. For each sample the mean $\bar{x}$ is calculated with an associated 95 % confidence interval.
- Confidence level 95 % for the parameter µ means that on average 95 of the confidence intervals calculated for $\bar{x}$ contain the the true parameter value μ (see Figure 31).
- This clearly means that on average 95 of the confidence intervals calculated for $\bar{x}$ need to overlap in some areas, so that the statement may be true in principle.
- **In practice,** only one sample is drawn, so that there is a risk of 5 % to randomly draw a sample, whose calculated confidence interval does not contain the true parameter μ.

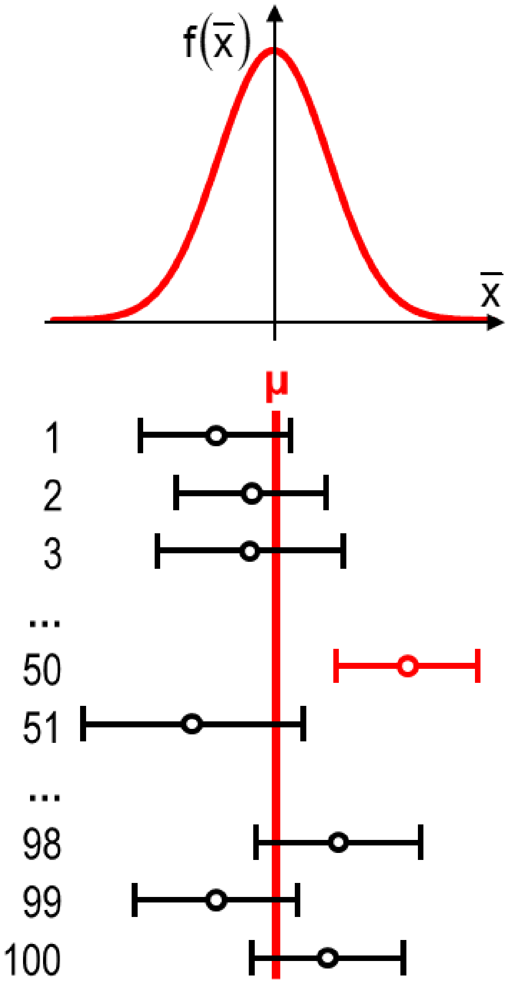


Figure 31: Confidence intervals

Calculation in the case of normally distributed data:

In a sample consisting of $n \geq 30$ normally distributed readings x with mean $\bar{x}$ and standard deviation s the true value for the standardized quantity $z = \frac{x - \bar{x}}{s}$ is to be expected with probability $(1 - \alpha) \cdot 100\ \%$ between the quantiles $u_{\alpha/2}$ and $u_{1-\alpha/2}$ of the standard normal distribution:

$$u_{\alpha/2} \leq \frac{x - \bar{x}}{s} \leq u_{1-\alpha/2} \qquad \text{or dissoved after } x \qquad \bar{x} + u_{\alpha/2} \cdot s \leq x \leq \bar{x} + u_{1-\alpha/2} \cdot s$$

Accordingly, the interval $\left[\bar{x} + u_{\alpha/2} \cdot s; u_{1-\alpha/2} \cdot s\right]$ is called a confidence interval for x on confidence level $(1 - \alpha) \cdot 100\ \%$.

The procedure for other distributions is analogous (e. g. t-distribution for positions with sample size $n < 30$, χ^2 distribution of variances). In contrast to the standard normal distribution, for these and numerous other distributions the sample size n in addition to the confidence level affects the confidence limits and so on the statistical parameters such as C_p (see for example, Appendix I.1, Figure 41, "trumpet curve").

C.4 Notes for selection of distribution models

The reliability of the results of a capability analysis is decisively dependent on the appropriate choice of the statistical distribution model for the **population**. Samples are merely snapshots. The conclusion from the distribution model of the sample to the distribution model of the population is therefore already problematic. The problem becomes further problematic with the trend towards less samples and a smaller sample size, that is, an increasingly inadequate statistical data base. An adequate distribution model of the population can often not be directly determined in this way and expertise is absolutely essential.

Software selects a statistical distribution model for a sample by means of statistical tests or regression coefficients. The selection criterion is of pure mathematical and not technical nature, meaning, the best possible assignment, for example as on an underlying histogram. A reasonable decision for or against a distribution model can be made exclusively on the basis of the technical realities and circumstances. This can so-called best-fit tests basically not afford, meaning, software can only support experts in selecting an appropriate distribution.

- The **extended normal distribution (END)** typically describes processes well, in which the expected value of an otherwise stable normal distribution gradually shifts (e. g., due to tool wear). Therefore, the END is a simple, but statistically relatively representative distribution model for unstable processes. Skewed and kurtotic distributions can in principle not be described perfectly by the END.
- The **mixed distribution (MD)** enables good adaptation to any measurement data distribution due to its mathematical conception. Compared to other distributions, it therefore often produces better capability indices, which are however technically not founded.

In spite of this, there is a rapidly increasing trend that has been observed to fully automate capability analyses, meaning, to relinquish it to software. The consequence is that results are hardly questioned afterwards. Especially when selecting statistical distribution models, it is not ensured that technical constraints are adequately addressed. Capability indices resulting in this way are consequently meaningless. This is especially critical in processes which are classified as unstable.

The automated selection of the distribution cannot effectively deal with this problem. Instead, in individual cases, analysis is necessary by qualified experts sufficiently familiar with the process, whether and to what extent the delivered result is compatible with the actual technical conditions and is relevant. The following basic factors have to be considered:

- A measurable characteristic (process result) is generally a superposition of (often not measurable individually) characteristics, whose respective expression is determined by the individual process steps that produce the characteristic. The statistical distribution of the process result is therefore also a superposition of the statistical distributions of the partial and/or interim results of the individual process steps. If these are primarily normally distributed, a normally distributed population for the process result is to be expected (central limit theorem of statistics).
- Besides there are characteristics that cannot be normally distributed in principal due to their definition, e. g., characteristics with natural lower limit 0 (see Appendix K) or characteristics with trend).

It is recommended to perform this analysis in advance, but no later than during the process start-up phase on a possibly broad and stable data base.

Not the distribution of the current sample but the distribution of the population is crucial.

qs-STAT® offers the possibility to select deviating corresponding distributions of the auto-detected distribution.

NOTE: Deviations from the pre-settings of the evaluation strategy must be justified and documented

C.5 Options for very large data sets

Chapter 8.1 "Estimate Values" already explained that the realistic estimate of the 0.135% and 99.865% quantiles of the data distribution for the determination of resilient capability indices are essential. This means that primarily the edge regions ("tails") of the distribution are relevant, which means, the best possible estimate of the probabilities of the occurrence of the minimum and maximum values.

The relatively small data sets from a statistical perspective that are usually available (50 values for C_{mk}, 125 values for C_{pk}) usually do not contain a sufficient number of values in the boundary areas in order to reliably determine the quantile required. The probabilities of very small and very large values must therefore be estimated with the help of a distribution function, which is determined on the basis of the frequently occurring values in the middle area around the expected value.

With increasing size of the data set, the number of available values increases in the peripheral areas, so that other options may arise in determining the required quantile.

- With sufficiently large data sets[14] the smallest and largest value among others can be used directly as an estimator for the 0.135% or 99.865% quantile without adjustment of a distribution model. However, an estimate based on only two measurements in the extreme location can naturally be extremely sensitive in regards to outliers and therefore **not** recommended.

- A significantly more efficient and more reliable variant of this method is to evaluate data from the two edge regions of the probability plot, i. e. the minimum and maximum values. *Experience has shown* (see notes below), that for this purpose, 25 - 30 data respectively are required, which often prove to be approximately normally distributed, meaning, building a straight line in the probability plot of the standard normal distribution. In this case, a regression line can be respectively adapted and determined to the y-value -3s or + 3s of the respective x value (0.135% - or 99.865% quantile). Data closest to the edge often show greater deviations from the normal distribution, and should not be included in the regression in this case. *Experience shows* that it is usually enough, to exclude 1 - 5 data.

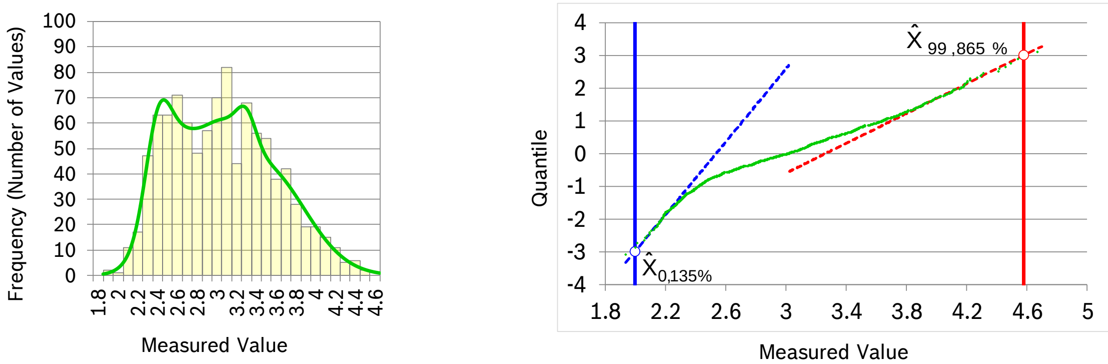


Figure 32: Determination of process variation by regression in the marginal areas

NOTE 1: The numbers listed are not to be understood as a requirement or recommendation, but as purely empirical values that have exclusively been proven under the boundary conditions performed previously and known investigations. Basically, it has to be assessed for each scenario, if these numbers are meaningful. Up to now there are no standard guidelines of any kind in this regard.

NOTE 2: It is expressly pointed out, that the choice and number of data and their compatibility with the normal distribution generally have a strong effect on the slope of the regression line and thus on the determined position of the quantile and the resulting variation range.

- Another variant is to use the distribution model of the mixture distribution.

---

[14] Example in [ISO 22514-2], Section 6.1.4: Data set with $n \geq 1{,}000$ single values; corresponds to the previous methods $M1_{4,5}$ and $M1_{5,5}$ of the withdrawn [ISO 21747], also known as range methods.

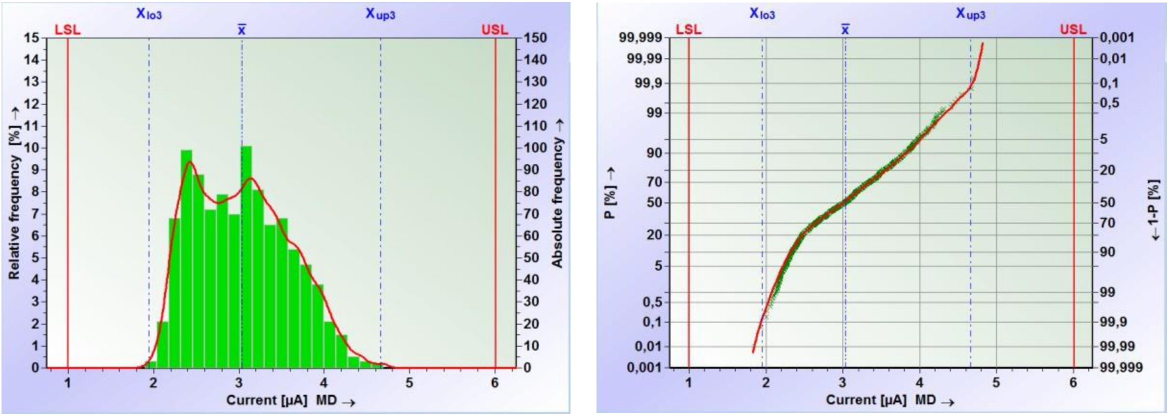


Figure 33: Determination of process variation by mixture distribution

The same data set ($n = 1{,}000$ individual values, limits $LL = 1\ \mu A$ and $UL = 6\ \mu A$) evaluated with this method returns the following results:

	Evaluation of extreme values	Evaluation of peripheral areas	Evaluation mixture distribution
$X_{0.135\ \%}$	1.94	2.00	1.95
$X_{99.865\ \%}$	4.70	4.58	4.66
$X_{99.865\ \%} - X_{0.135\ \%}$	2.76	2.58	2.71
C_p	1.81	1.94	1.84
C_{pk}	1.77	1.91	1.81

Table 3: Evaluation results for a large data set

In this particular example the results for C_p vary in the range 1.81 to 1.94. The results for C_{pk} behave similarly.

- It is plausible that the evaluation of the extreme values tend to the worst results of the three provided variants, since the extreme values of the dataset are frequently below and above the actual 0.135% - or 99.865% quantile and thus lead to greater than the actual variation range.
- A similar trend *can* — such as in the present case — be observed in the mixed distribution. This is due to the mathematical concept of the mixed distribution, which adapts to the database almost arbitrarily, so that in particular in this model, a few data points can significantly affect the outcome close to the edge. However, the result often also depends on the software used, because several, sometimes very different algorithms exist for the adaption. For example, the decomposition by means of so-called statistical moments[15] (for example, in analysis with qs-STAT®) will lead to slightly different results than the simple decomposition into multiple, overlapping normal distributions (as with manual analysis). In the latter case, a tendency towards the results which are to be expected in the evaluation of the data in the edge region is plausible, because the outermost normally distributed partial collective in the total distribution often also determines the extreme edges of the overall distribution.
- The evaluation of the two edge areas often leads to the best results. But this depends very sensitively on the selection of data included in the analysis, (see above Note 2) which are correspondingly easy to manipulate and may even violate statistical principles.

[15] Standard deviation, skewness and kurtosis are the statistical moments of the 2nd to 4th order.

D Statistical distribution tests

The evaluation configurations implemented in qs-STAT® for sample and process analysis contain some statistical tests, which analyze the data set to be evaluated with regard to several statistical questions and thus enable an automatic distribution assignment. Please refer to [Booklet 3] for the basic procedure of a statistical test.

In this Appendix D some examples are explicitly listed. Some of the classical test procedures are only suitable for certain sample sizes. In recent years, however, there have been improvements or new variants of the tests.

The aim of the evaluation configurations is to achieve the best possible, realistic adaptation of the theoretical distribution model to the empirical data set.

D.1 Tests for normal distribution

There are numerous comparative studies of normality tests with regard to their test strength, e. g. [Seier]. If the null hypothesis is maintained on the basis of the test result, although there is actually a deviation from the normal distribution, it is an error of the second kind (type 2 error, [Booklet 3]). The test strength indicates the probability of avoiding a type 2 error. The test strength depends on the type of deviation and the sample size. You can only increase the test strength of a test by selecting a larger sample size. For this reason, different tests are used depending on the sample size:

Test	**Range of application**
Shapiro-Wilk test	$8 \leq n \leq 50$
Epps-Pulley test	$51 \leq n \leq 200$
Asymmetry	$n \geq 201$
Kurtosis	$n \geq 201$

D.1.1 Normality test using skewness and kurtosis

The most common statistical tests for normal distribution include tests of symmetry and form of data distribution. The (empirical) **Skewness** $\beta_1 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^3$ is a measure of the direction and intensity of a distribution's asymmetry:

- $\beta_1 < 0$ left-skewed (or rather right-steep) distribution;
- $\beta_1 = 0$ Normal distribution (no asymmetry);
- $\beta_1 > 0$ right-skewed (or rather left-steep) distribution.

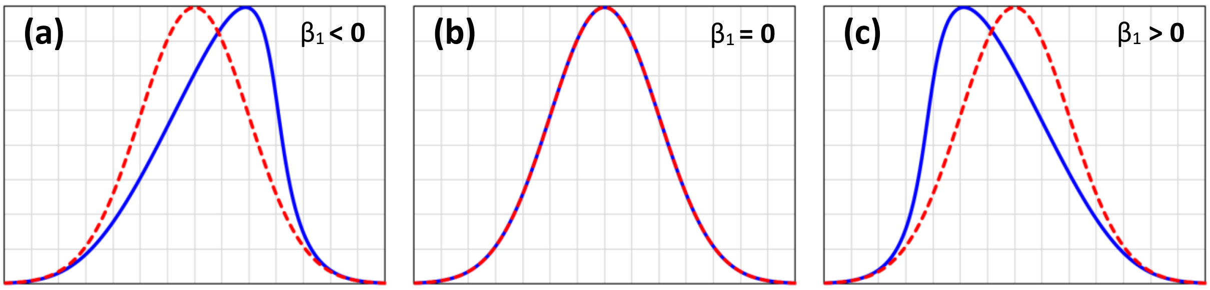


Figure 34: (a) left-skewed, (b) symmetric and (c) right-skewed distribution

With the help of a statistical test (see e. g. [Booklet 3]), it is determined whether $\beta_1 = 0$ the confidence interval of the determined β_1 value is located at a certain level of confidence (typically 95 %). If this is the case, the measurement data follows, with this degree of confidence, normal distribution.

The (empirical) **Kurtosis** $\beta_2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^4$ is a measure of the flatness or slope of a distribution compared with the normal distribution:

- $\beta_2 < 3$ Distribution has „more broad-peak" than the normal distribution;
- $\beta_2 = 3$ normal distribution;
- $\beta_2 > 3$ Distribution has „more slender peak" than the normal distribution.

Instead of kurtosis β_2 the so-called excess $\gamma_2 = \beta_2 - 3$ is used, at which the zero point $\gamma_2 = 0$ is defined through the normal distribution.

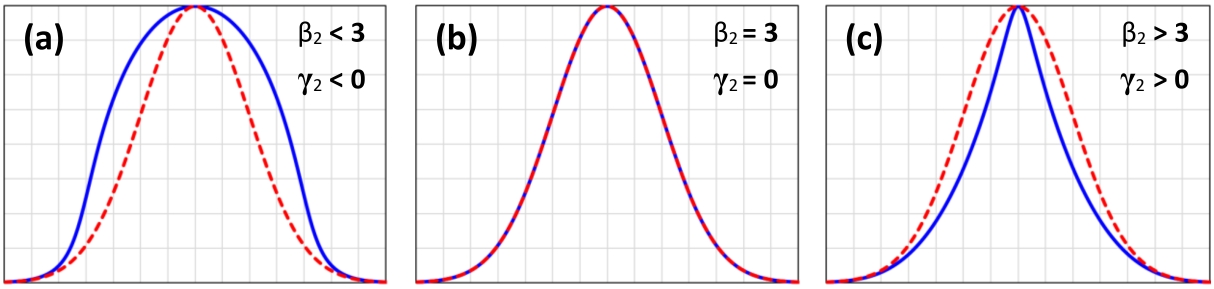


Figure 35: (a) platykurtic, (b) normal and (c) leptokurtic distribution

With the help of a statistical test (see for example [Booklet 3]), it is determined whether $\beta_2 = 3$ or $\gamma_2 = 0$ at a certain level of confidence (typically 95%). If this is the case, the measurement data will follow, with this degree of confidence, normal distribution.

D.1.2 Shapiro-Wilk test

According to [ISO 5479], the Shapiro-Wilk normality test is suitable for small sample sizes with $8 \leq n \leq 50$. It uses the variance s^2 and the sum of weighted differences of the individual values ordered in ascending order to calculate the test statistic $W = \frac{b^2}{(n-1) \cdot s^2}$

$x_1 \leq x_2 \leq x_i, \ldots, x_{n-1} \leq x_n.$

For even n

$$b = \sum_{i=1}^{n/2} a_{n-i+1} \cdot (x_{n-i+1} - x_i) = a_n \cdot (x_n - x_1) + \ldots + a_{n/2} \cdot \left(x_{n/2+1} - x_{n/2}\right).$$

With odd n

$$b = \sum_{i=1}^{(n-1)/2} a_{n-i+1} \cdot (x_{n-i+1} - x_i) = a_n \cdot (x_n - x_1) + \ldots + a_{(n-1)/2} \cdot \left(x_{(n-1)/2} - x_{(n+1)/2+1}\right).$$

The coefficients (weighting factors) a_i are tabulated in [ISO 5479] and [Wilrich], for instance.

Note: In case of odd n, the median is not used. The a_i become smaller with increasing i, i. e., a_n is the largest coefficient; the difference of the extreme values $x_n - x_1$ is taken into account most strongly.

According to [Seier], the test is sensitive to too coarse a rounding of the measured values, i. e. if the place value is greater than 10 % of the standard deviation. See also Section 6.3.

D.1.3 Epps-Pulley test

The Epps-Pulley normality test uses the test statistic T_{EP} based on the characteristic functions of the sample and the normal distribution.

$$T_{EP} = \frac{2}{n} \cdot \sum_{k=2}^{n} \sum_{j=1}^{k-1} e^{-\frac{1}{2} \cdot \frac{\left(x_j - x_k\right)^2}{m_2}} - \sqrt{2} \cdot \sum_{j=1}^{n} e^{-\frac{1}{4} \cdot \frac{\left(x_j - \bar{x}\right)^2}{m_2}} + \frac{n}{\sqrt{3}} + 1$$

$m_2 = \frac{n-1}{n} \cdot s^2$ is the 2nd central moment and s^2 the variance of the sample.

One of the advantages of the test is that it does not require tabulated constants. [ISO 5479] explicitly underlines the high test strength of this test and specifies the 95 % and 99 % quantiles of the test statistic for sample sizes in the range $51 < n \leq 200$. [Epps], [ISO 5479], [Schulze]

D.1.4 Test for instantaneous normal distribution (extended Shapiro-Wilk test)

Often the size n of the individual samples is quite small, e. g. $n = 3$ or $n = 5$. Therefore, a normality test applied to each individual sample could not detect larger deviations from the normal distribution. If the individual samples show variation of the means, it also does not make sense to combine all individual samples into a total sample with $m \cdot n$ values.

The extended (modified) Shapiro Wilk test assumes that a data set of m samples with a sample size of n is available. The test checks whether the independent individual samples can originate from a normally distributed population (null hypothesis) or not. The test is quite computer-intensive because a test quantity W_j with $j = 1, 2, \ldots, m$ is calculated for each individual sample using tabulated coefficients. The W_j are finally combined to a total test statistic. If the null hypothesis is correct, the total test statistic is approximately standard normally distributed.

The b_j are calculated as in D.1.2, also $W_j = \frac{b_j^2}{(n-1) \cdot s_j^2}$. Finally,

$$W = \sqrt{m} \cdot \gamma(n) + \frac{\delta(n)}{\sqrt{m}} \cdot \sum_{j=1}^{m} ln\left(\frac{W_j - \varepsilon(n)}{1 - W_j}\right)$$ with tabulated coefficients $\gamma(n), \delta(n)$ and $\varepsilon(n)$.

The null hypothesis is rejected if $W < -u_{1-\alpha}$. Where $u_{1-\alpha}$ is the quantile of the standard normal distribution for the error probability α, e. g. $-u_{1-0.05} \approx -1.645$.

Lit.: [ISO 5479], [Schulze]

D.2 Tests on arbitrary distributions (except ND)

The Chi-squared test (χ^2-test) can be used to check whether the data set at hand is compatible with any given distribution. To do this, the empirical data must be available in classified form, e. g. as a histogram with absolute frequencies over a class division.

Simply put, the test assesses how much the columns of the histogram deviate up or down from the theoretical frequencies in the corresponding classes. The determined test statistic is compared with a quantile of the χ^2 distribution.

k is the number of classes, b_i the observed frequency in class i. e_i is the expected (theoretical) frequency in class i if the data come from the assumed distribution F_0, i. e. if H_0 is correct.

The total number of values is $n = \sum_{i=1}^{k} b_i$. The expected frequency can be calculated according to $e_i = n * p_i$, where $p_i = F_0\left(x_{i;\,ob}\right) - F_0\left(x_{i;\,un}\right)$ can be determined from the values of the distribution function at the upper limit $x_{i;\,ob}$ and lower limit $x_{i;\,un}$ of the i-th class. The test statistic is then

$$\chi_{test}^2 = \sum_{i=1}^{k} \frac{(b_i - e_i)^2}{e_i}.$$

The null hypothesis is rejected if $\chi_{test}^2 > \chi_{k-1;\,1-\alpha}^2$.

Since e_i is in the denominator, it must not become zero. According to [Hartung] and [Wilrich], no value e_i may be smaller than 1 and no more than 20 % of e_i may be less than 5 (see also [DIN EN 61710]).

> *Note: The Chi-squared test requires classified data. The classification can have an influence on the test result.*

The tests mentioned under D.5 are preferred as normality tests.

E Distribution models according to ISO 22514-2

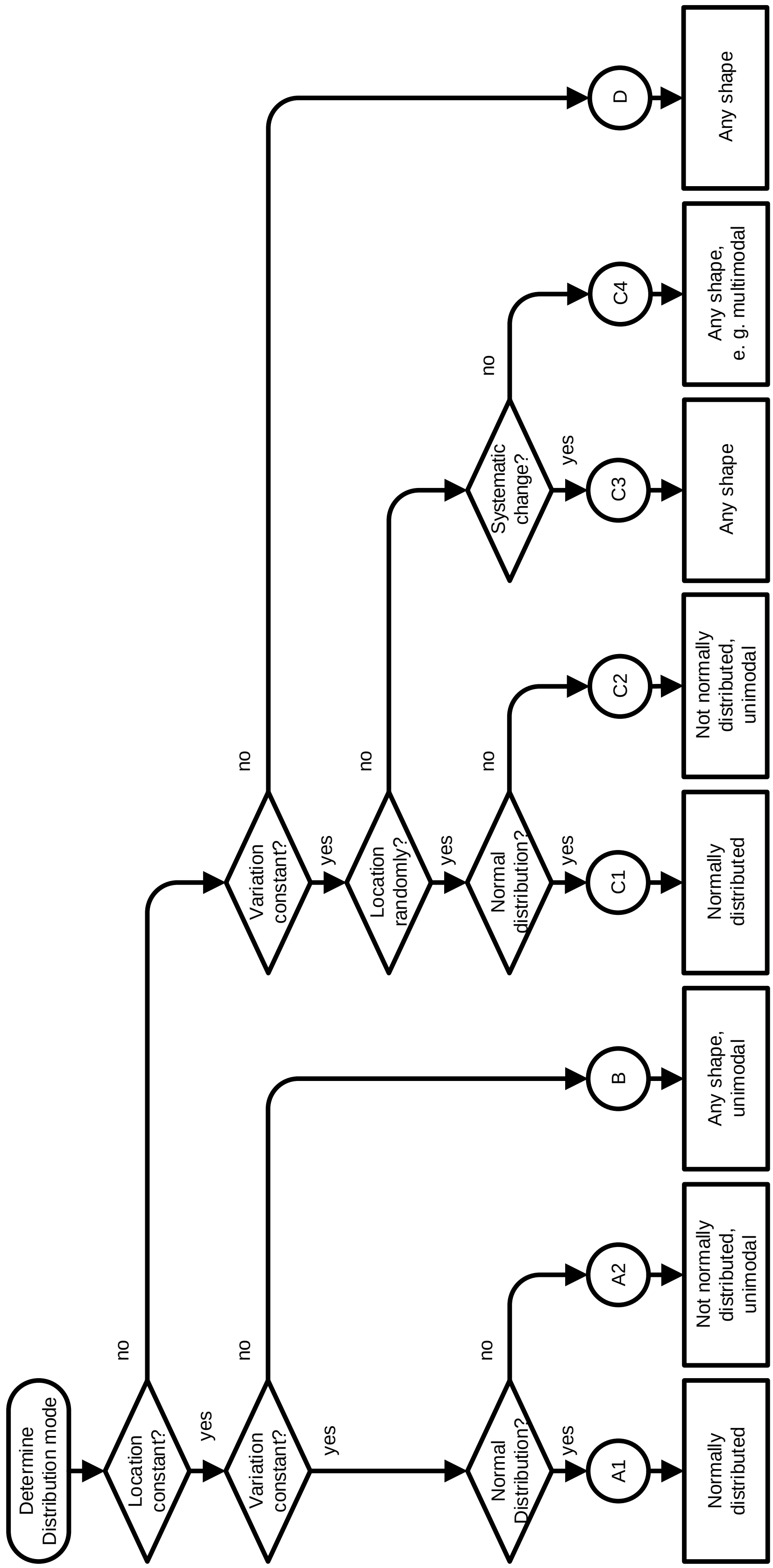


Figure 36: Distribution models according to ISO 22514-2

Estimator		**Designation according to ISO 22514-2**	**Distribution model acc. to ISO 22514-2**							
			Stable location			Location not stable				
Location	Variation		A1	A2	B	C1	C2	C3	C4	D
$\bar{x}$	$X_{99.865\,\%} - X_{0.135\,\%}$	$M_{1,1}$	✓		✓					
	$\sqrt{s^2}$	$M_{1,2}$	✓							
	$\frac{\bar{s}}{c_4}$	$M_{1,3}$	✓							
	$\frac{\bar{R}}{d_2}$	$M_{1,4}$	✓							
	s_{total}	$M_{1,5}$	✓		✓					
$\tilde{x}$	$X_{99.865\,\%} - X_{0.135\,\%}$	$M_{2,1}$	✓	✓	✓	✓	✓	✓	✓	✓
	$\sqrt{s^2}$	$M_{2,2}$	✓							
	$\frac{\bar{s}}{c_4}$	$M_{2,3}$	✓							
	$\frac{\bar{R}}{d_2}$	$M_{2,4}$	✓							
	s_{total}	$M_{2,5}$	✓	✓	✓	✓				✓
$\bar{\bar{x}}$	$X_{99.865\,\%} - X_{0.135\,\%}$	$M_{3,1}$	✓							
	$\sqrt{s^2}$	$M_{3,2}$	✓							
	$\frac{\bar{s}}{c_4}$	$M_{3,3}$	✓							
	$\frac{\bar{R}}{d_2}$	$M_{3,4}$	✓							
	s_{total}	$M_{3,5}$	✓							
$\bar{\tilde{x}}$	$X_{99.865\,\%} - X_{0.135\,\%}$	$M_{4,1}$	✓	✓	✓					
	$\sqrt{s^2}$	$M_{4,2}$	✓							
	$\frac{\bar{s}}{c_4}$	$M_{4,3}$	✓							
	$\frac{\bar{R}}{d_2}$	$M_{4,4}$	✓							
	s_{total}	$M_{4,5}$	✓	✓	✓					

NOTE: The methods M1,1 and M1,5 are applicable to the distribution model B, but M3,1 and M3,5 are not.
Both groups of methods differ only in the method of counting the data values, but not in the result.
The different applicability according to the standard is therefore not plausible.

Table 4: Applicability of the calculation methods on distribution models according to ISO 22514-2

F Impact of the measurement process variation

Both the manufacturing process and the measurement process with which the production result is checked possess variation. The variation of the measurement results is therefore caused proportionately by **both** processes. Based on the measurement results, however, only the manufacturing process should be assessed. This is only possible if the variation of the measurement process is sufficiently small in comparison to the manufacturing process.

The variation range $6 \cdot \sigma$ of a manufacturing process is often — especially for normally distributed characteristic values — estimated based on the empirical standard deviation s of the measurement results: $\sigma \approx \hat{\sigma} = s$. The notation $\hat{\sigma}$ denotes an estimator to which s is assigned as an estimate. The standard deviation s is made up of the parts s_P (caused by the **p**roduction process) and s_M by the **m**easurement process) together:

$$s = \sqrt{s_P^2 + s_M^2}\,. \tag{F.1}$$

For normally distributed characteristic values, the influence of the measurement process variation using the calculation rule for the potential capability index C_p can be quite easily described mathematically. Eq. (F.1) used and the term T/6 drawn into the root:

$$C_p = \frac{T}{6 \cdot s} = \frac{T}{6 \cdot \sqrt{s_P^2 + s_M^2}} = \frac{1}{\sqrt{\left(\frac{6 \cdot s_P}{T}\right)^2 + \left(\frac{6 \cdot s_M}{T}\right)^2}} \tag{F.2}$$

The first summand of the root can be interpreted as the inverse ***of the actual*** manufacturing process capability C_p^*, which is free from the influence of the measurement process variation:

$$C_p^* = \frac{T}{6 \cdot s_P}. \tag{F.3}$$

The standard deviation s_M of the measurement process is commonly designated as GRR ($s_M = GRR$) and expressed as a percentage of a reference value in the measurement system analysis. With respect to the characteristic tolerance T (see [Booklet 10]), the following applies:

$$\%GRR = \frac{6 \cdot GRR}{T} \cdot 100\,\% = \frac{6 \cdot s_M}{T} \cdot 100\,\% \tag{F.4}$$

The Eq. (F.3) and (F.4) used in Gl. (F.2) equals

$$C_p = \frac{1}{\sqrt{\left(\frac{1}{C_p^*}\right)^2 + \left(\frac{\%GRR}{100\,\%}\right)^2}} \tag{F.5}$$

The **observed** process capability C_p depending on the actual process capability C_p^* is often depicted as an array of curves with parameter $\%GRR$ (Figure 37).

NOTES:

- *It is expressly pointed out that it is **not** permissible, to “optimize” calculated C_p values using equation (F.5) or Figure 36 and specify C_p^* instead of C_p.*
- *The statements in this section merely serve to develop an understanding of the fundamental relationships and a sense of magnitudes of deviations such as $C_p - C_p^*$ and characteristics such as % GRR when e. g. given maximum deviations $C_p - C_p^*$ must be maintained.*

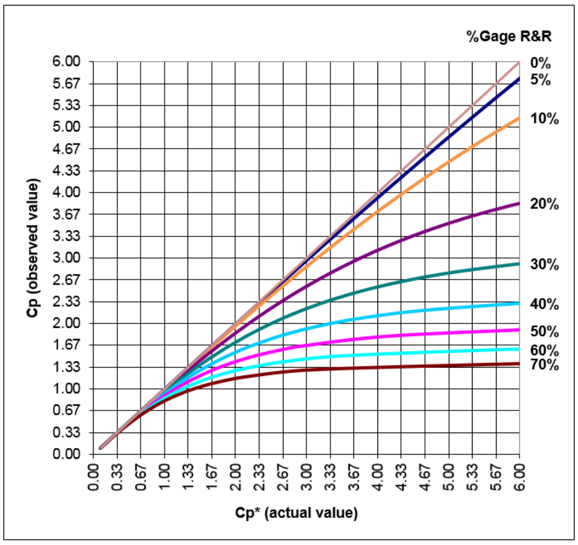


Figure 37: Relationship between observed and actual process capability

Thereafter, in an actual (i. e. theoretically ideal) manufacturing process capability $C_p^* = 2.67$ ($\%GRR = 0\ \%$) and measurement process variation according to Figure 38(a) the manufacturing process variation according to Figure 38(b) with the capabilities

$C_p = 2.58$ ($\%GRR = 10\ \%$),

$C_p = 2.08$ ($\%GRR = 30\ \%$) and

$C_p = 1.60$ ($\%GRR = 50\ \%$)

are observed.

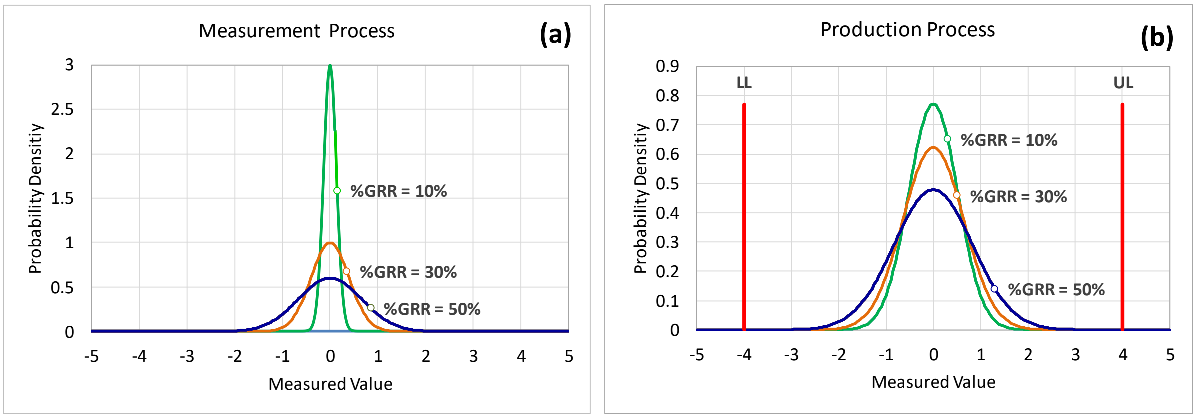


Figure 38: Impact of the measurement process variation on observed manufacturing process variation

The key finding from Eq. (F.5) — or more clearly Figure 37 — is, that in particular higher C_p values only represent a meaningful result if the measurement process variation is comparatively low. This applies — regardless of Eq. (F.5) — also in the general case of not normally distributed data. In the specific case, according to Eq. (F.5), the relationship $\left(\frac{1}{C_p^*}\right)^2 \gg \left(\frac{\%GRR}{100\ \%}\right)^2$ must be adequately met. This means that for a given $\%GRR$, a range of values exists for $\frac{1}{C_p^*}$, which must not be undershot, meaning C_p^* must not become too large.

EXAMPLE: In an actual capability of von $C_p^* = 6.0$ and $\%GRR = 30\ \%$, the observed capability just under $C_p \approx 3.0$ is only about 50% of the actual capability (see Figure 37). Even at $C_p^* = 2.0$, at $C_p \approx 1.67$ an almost 17% smaller value is still observed. On the other hand e. g. to comply with the maximum difference $C_p - C_p^* \leq 0.1$, $\%GRR \approx 19{,}8\ \%$ at $C_p = 1.33$ is required and $\%GRR \approx 2.1\ \%$ at $C_p = 6.0$.

G Stable processes and processes in control

The distinction of stable and controlled processes often leads to misunderstandings due to inconsistent contents of different standards. Causes here are i. a. difficulties in translating from English into the national language.

ISO 3534-2

The English edition of the standard [ISO 3534-2] from 2006 use terms "stable process" and "process in a state of statistical control" interchangeably. Note 1 analogously supplements, that such a process behaves, *as though the samples from the process at any time are simple random samples from the same population*. Note 4 explicitly states that processes with increasing change of mean and/or standard deviation, as e. g. such as occurs due to wear-out of tools, are also subject to systematic variation causes, which are not considered as a result of random causes, and *random samples from such processes cannot originate from the same population* (see Chapter "*Terms and Definitions"* or [ISO 3534-2, 2.2.7]).

ISO 21747 (withdrawn)

Terms, definitions and comments are taken over unchanged from the English edition of the standard [ISO 3534-2] in the English edition of the now invalid standard [ISO 21747].

ISO 22514-1

Terms and definitions of the English edition of the standard [ISO 3534-2] will be finally adopted in the English edition of the standard [ISO 22514-1], but only with Note 1 and 2, which are now recognized as Notes 2 and 3.

The omitted Note 4 became to a certain extent independent in the standard [ISO 22514-1] and used accordingly to define the newly introduced concept of "product characteristic in control": "*Product characteristic parameter of the distribution of the characteristic values of which practically do not change or do change only in a known manner or within known limits"*.[16, 17]

[16] The English expressions "*in a state of statistical control"* and "*in control"* are often considered as equivalent by native English speakers.

[17] The German edition points out inconsistencies in the standards in a national footnote.

Understanding at Bosch

The definitions of the new standard [ISO 22514-1] are very close or complimentary to the definitions according [DIN 55350-11], and are generally congruent with the previous understanding at Bosch, which will therefore continue unchanged:

- **Stable process:** Process subject only to random causes *([ISO 22514-1, 3.1.21])*

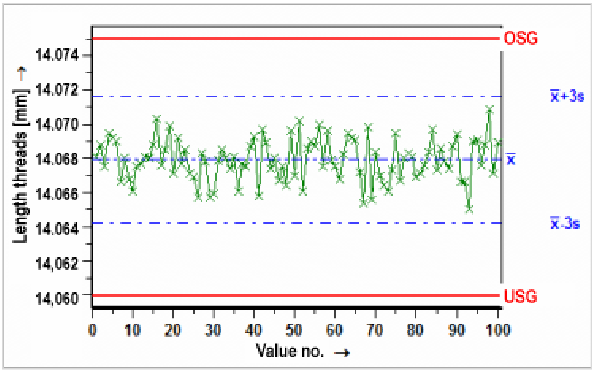


Figure 39: Example of a stable process

NOTE 1: For stable processes, only random mean changes occur, which cancel each other out to the mean.

- **Controlled process (process in control):** Unstable process, in which parameter of the distribution of the characteristic values of which practically do not change or do change only in a known manner or within known limits (with reference to [ISO 22514-1, 3.1.20] and [DIN 55350-11, 3.11.1 & 3.11.2])

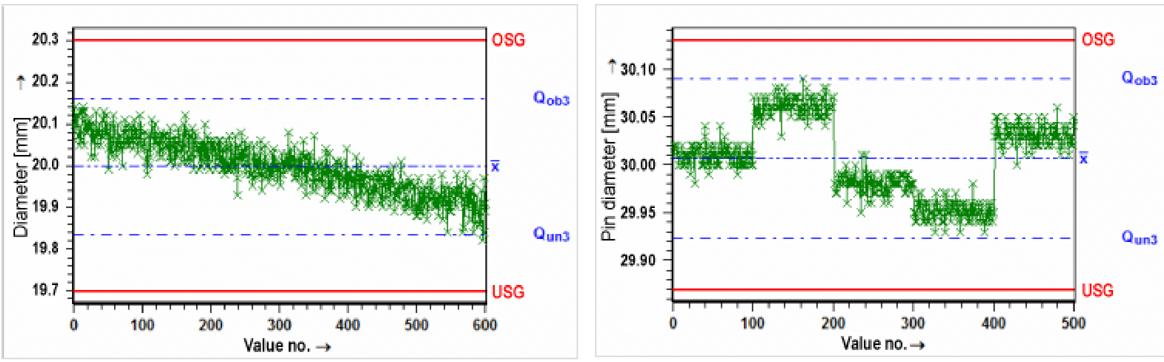


Figure 40: Examples of unstable but controlled processes

NOTE 2: In unstable but controlled processes, systematic changes of the mean occur, which are within certain limits and whose causes are known (for example, tool wear).

H Definition of $C_{p(k)}$ and $P_{p(k)}$ acc. to ISO 22514 and AIAG SPC

The definitions of the process capability indices C_p and C_{pk} (capability) and process performance indices P_p and P_{pk} (performance) according to [ISO 22514-2] and [AIAG SPC] are different.

Both approaches assume that at certain time intervals, samples are drawn from the total production volume. Positions and variations of each sample overlap to a total distribution (see Figure 41).

NOTE 1: Based on the properties of the individual samples and their temporal behavior, [ISO 22514-2] differentiates eight different models for the overall distribution, of which only two models are normal distributions (see Appendix E). While [ISO 22514-2] allows any distribution of measurements for the most part, the approach according to [AIAG SPC] based on sufficiently normally distributed values. In the case of skewed distributions, normally distributed values can often be produced by appropriate transformation.

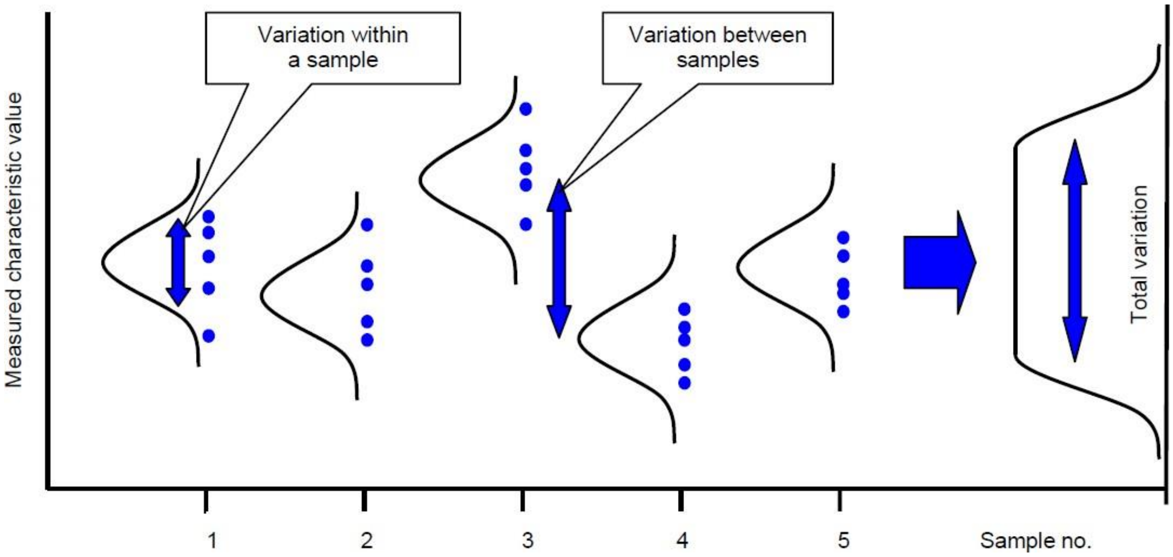


Figure 41: Contribution of variation and position of the individual samples to the total variation

ISO 22514

According to [ISO 22514-2], C_p and P_p are calculated by using the same calculation rule; the same applies to C_{pk} and P_{pk}:

$$\begin{matrix} C_p \\ P_p \end{matrix} = \frac{UL-LL}{X_{99,865\,\%}-X_{0,135\,\%}}; \qquad \begin{matrix} C_{pk} \\ P_{pk} \end{matrix} = min\left(\frac{X_{mid}-LL}{X_{mid}-X_{0,135\,\%}};\ \frac{UL-X_{mid}}{X_{99,865\,\%}-X_{mid}}\right)$$

Whether the calculated results from the measurements are interpreted as process capability or process performance and accordingly are denoted by C_p and C_{pk} or P_p and P_{pk}, is decided exclusively by the process stability. Stable processes are associated with C_p and C_{pk} unstable processes P_p and P_{pk}. Unstable are particularly processes with significant location differences between individual samples, i. e. significant variation of sample means. In addition, other criteria such as the variation of the sample variances and statistically unlikely behavior (Run, Trend, Middle Third) are possibly relevant.

*NOTE 2: In contrast to the approach according to [AIAG SPC], the calculation results always contain in principle both variation parts, that is, the variation within **and** between samples. The analysis of the process stability assesses only if the variation component between samples (relative to the variation component within the samples) should be evaluated as significant, but without changing the numerical calculation result. To that extent, the result may be i. a. less favorable than according to [AIAG SPC].*

AIAG SPC

In contrast to [ISO 22514-2] defines [AIAG SPC] indices C_p and C_{pk} exclusively by the variation **within** the individual samples, i. e. there are no variations between samples included:

$$C_p = \frac{UL - LL}{6 \cdot \sigma_C}; \qquad \text{and} \qquad C_{pk} = min\left(\frac{X_{mid} - LL}{3 \cdot \sigma_C}; \frac{UL - X_{mid}}{3 \cdot \sigma_C}\right)$$

with

$$\sigma_C \approx \hat{\sigma}_C = \sqrt{\overline{s^2}} = \sqrt{\frac{1}{m} \cdot \sum_{k=1}^{m} s_k^2} = \sqrt{\frac{1}{m} \cdot \sum_{k=1}^{m} \left\{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_{ik} - \bar{x}_k)^2\right\}} \qquad \bar{x}_k = \frac{1}{n} \cdot \sum_{i=1}^{n} x_{ik}$$

and

σ_C	Standard deviation of the population
$\hat{\sigma}_C$	Estimate for the standard deviation σ_C of the population
m	Number of samples
n	Number of measurements per sample
k	Number of the sample: $1 \le k \le m$
i	Number of the measurement within a sample: $1 \le i \le n$
x_{ik}	Measurement no. i in sample no. k
$\bar{x}_k$	Mean of the measurements x_{ik} in sample no. k
s_k	Standard deviation of the measurements x_{ik} in sample no. k

Since the estimator $\hat{\sigma}_C$ is determined solely from the empirical sample variances s_k^2, it contains no variation components between the individual samples. This means that effects such as e. g. drift through tool wear can **not** be detected, which will certainly change the position of the sample, but not necessarily the mean variation. For this reason, are C_p und C_{pk} understood as **short-term capability** (within the context of [AIAG SPC]).

P_p and P_{pk} however, are defined over the total variation of all measurements x_{ik} which additionally also contain the variation components **between** the individual samples:

$$P_p = \frac{UL - LL}{6 \cdot \sigma_P}; \qquad \text{and} \qquad P_{pk} = min\left(\frac{X_{mid} - LL}{3 \cdot \sigma_P}; \frac{UL - X_{mid}}{3 \cdot \sigma_P}\right)$$

with

$$\sigma_P \approx \hat{\sigma}_P = \sqrt{\frac{1}{N-1} \cdot \sum_{j=1}^{N} \left(x_j - \bar{x}\right)^2} = \sqrt{\frac{1}{m \cdot n - 1} \cdot \sum_{k=1}^{m} \sum_{i=1}^{n} (x_{ik} - \bar{x})^2} \tag{H.1}$$

$$\bar{x} = \frac{1}{N} \cdot \sum_{j=1}^{N} x_j = \frac{1}{m \cdot n} \cdot \sum_{k=1}^{m} \sum_{i=1}^{n} x_{ik} = \frac{1}{m} \cdot \sum_{k=1}^{m} \bar{x}_k = \bar{\bar{x}}$$

and

σ_P Standard deviation of the population

$\hat{\sigma}_P$ Estimate for the standard deviation σ_P of the population

N Total of all measurements: $N = m \cdot n$

j Number of the measurement within the population of all measurements: $1 \leq j \leq N$; $j = (k-1) \cdot n + i$

x_j measurement no. j within the population of all measurements: $x_j = x_{(k-1) \cdot n+i} = x_{ik}$

$\bar{x}$ Mean value of the population of all measurements x_j

$\bar{\bar{x}}$ Mean of all sample means $\bar{x}_k$

By algebraic conversion of Eq. (H.1) $\hat{\sigma}_P$ according to

$$\hat{\sigma}_P = \sqrt{\frac{m \cdot (n-1)}{m \cdot n - 1} \cdot \overline{s^2} + \frac{(m-1) \cdot n}{m \cdot n - 1} \cdot s_{\bar{x}}^2} \quad \text{(H.2)}$$

can be splitted in the mean sample variance $\overline{s^2}$ (variation **within** the samples) and in the variance

$$s_{\bar{x}}^2 = \frac{1}{m-1} \cdot \sum_{k=1}^{m} (\bar{x}_k - \bar{\bar{x}})^2$$

of the sample means $\bar{x}_k$ (variation **between** samples). If the mean variation corresponds to the random variation

$$s_{\bar{x}} \approx \sqrt{\frac{\overline{s^2}}{n}}$$

which is expected based on the combined individual sampling readings in the mean, the Eq. will be reduced (H.2) to the form

$$\hat{\sigma}_P \approx \sqrt{\overline{s^2}}$$

i. e. the variation between samples is random and therefore insignificant. In this case the following applies

$$\hat{\sigma}_P \approx \hat{\sigma}_C$$

i. e. with sufficiently stable process location the results for C_p and P_p do not differ significantly and are statistically regarded as equal. The same applies for C_{pk} and P_{pk}.

NOTE 3: In practice, Pp and Ppk are usually smaller than Cp and Cpk. Therefore it is desirable to optimize the process, so that this difference is as small as possible.

Since the estimator $\hat{\sigma}_P$ also contains possible effects such as e. g. drift through tool wear, which will change the positions of the individual samples over time, the performance indices P_p and P_{pk} are also understood as **long-term capability** (within the meaning of [AIAG SPC]).

For stable processes, short- and long-term capability do not differ significantly.

I Procedure in case of insufficient number of parts

To evaluate the machine and process capability or process performance, a certain minimum number of parts are needed to ensure a sufficient “reliability” of the statistical statements. This minimum number of parts, i. e. measurements at 1 measurement/part, is frequently not attained in the case of very small manufacturing lots. In addition to small production volumes, reasons also include long production or measuring times per part, high unit costs, destructive tests.

While statistical methods in the range from $n' \geq 25$ parts with $n = n'$ readings are still applicable to a limited extent, the application in the area $n' < 25$ is increasingly questionable, as the “reliability” of the determined statistical results drastically decreases. Currently the following approaches, among others, are discussed and in use in practice:

1. Adjustment of the minimum requirements on capability and performance, depending upon the number of available measurements (see Appendix I.1);
2. Summary of the measurements of the same or comparable characteristics of the same or different parts (see Appendix I.2);
3. Use of the tolerance utilization %T as **non-statistical** acceptance criterion (see Appendix I.3) if neither approach according point 1 or 2 nor 100%-inspection is possible

Decisive for the applicability of the individual approaches and criteria for the qualification of production equipment and manufacturing processes are the available database and their composition:

<table>
<tr><th>Case</th><th>Number of parts n‘</th><th>Number of measurements n</th><th>Procedure</th><th>Acceptance criterion</th></tr>
<tr><td>A</td><td>$\boldsymbol{n' \geq 25}$
parts of the same part number</td><td rowspan="2">$\boldsymbol{n \geq 25}$
number of measurements equal to number of parts ($n' = n$)</td><td rowspan="3">Determining the required capability and performance indices:
C_m, C_{mk}
C_p, C_{pk}
P_p, P_{pk}
C_{p-ST}, C_{pk-ST}
P_{p-ST}, P_{pk-ST}</td><td rowspan="3">Compliance with the minimum levels of capability and performance indices *adapted to the available number of readings*

Example:
$C_{mk} \geq 1.88$ (n = 30)
$C_{pk} \geq 2.07$ (n = 35)</td></tr>
<tr><td>B</td><td>$\boldsymbol{n' \geq 25}$
By grouping parts with different part numbers, the required number of pieces $n' \geq 25$ will be reached</td></tr>
<tr><td>C</td><td rowspan="2">$\boldsymbol{n' < 25}$
Despite grouping parts with different part numbers, the required number of parts $n' \geq 25$ **is not** reached.</td><td>$\boldsymbol{n \geq 25}$
By grouping and scaling of measurements of various characteristics on each part, the required number of measurements $n \geq 25$ will be reached ($n' < n$)</td></tr>
<tr><td>D</td><td>**n < 25**
Despite grouping and scaling of measurements of various characteristics on each part, the required number of measurements $n \geq 25$ **will not** be reached</td><td>%T</td><td>All measurements within
75 % of $T - 2 \cdot U$</td></tr>
</table>

Table 5: Qualification of production equipment and processes with insufficient number of parts

I.1 Dynamization of the requirements on capability and performance

The “reliability” of a statistical statement is quantitatively determined by the confidence interval at a given confidence level $1-\alpha$. Clearly, that is the width of the interval in which the results of a statistical analysis with probability $1-\alpha$ is to be expected: the smaller the width, the more “reliable” the statistical result.

The width of the confidence interval at a given confidence level $1-\alpha$ is essentially determined by the number of individual components (for example, measurement results) from which the statistical results (for example, the variance) is calculated: the fewer the components (i. e., the smaller the sample size), the greater the width of the confidence interval, i. e. the more “less reliable” the statistical result.

The potential capability index $C_p = \frac{T}{6 \cdot s}$ depends solely on the empirical standard deviation s, which act as a estimator $\hat{\sigma}$ for the standard deviation σ of the population: for example, $\sigma \approx \hat{\sigma} = s$ in the simplest case or an estimator according to Chapter 8. The variation of s is described by the so-called χ^2 distribution. The confidence limits are calculated as the quantile of this distribution:

$$C_{p\,(min)} = \sqrt{\frac{\chi^2_{n-1;\,\alpha/2}}{n-1}} \cdot C_p \qquad C_{p\,(max)} = \sqrt{\frac{\chi^2_{n-1;\,1-\alpha/2}}{n-1}} \cdot C_p \qquad C_{p\,(min)} \le C_p \le C_{p\,(max)}$$

$C_{p\,(min)}$ and $C_{p\,(max)}$ represent the lower and upper confidence limit for C_p at n measurements and confidence level $1-\alpha$. These limits at a fixed confidence level $1-\alpha$ plotted against the number of measurements is often referred to as a “trumpet curve” presentation:

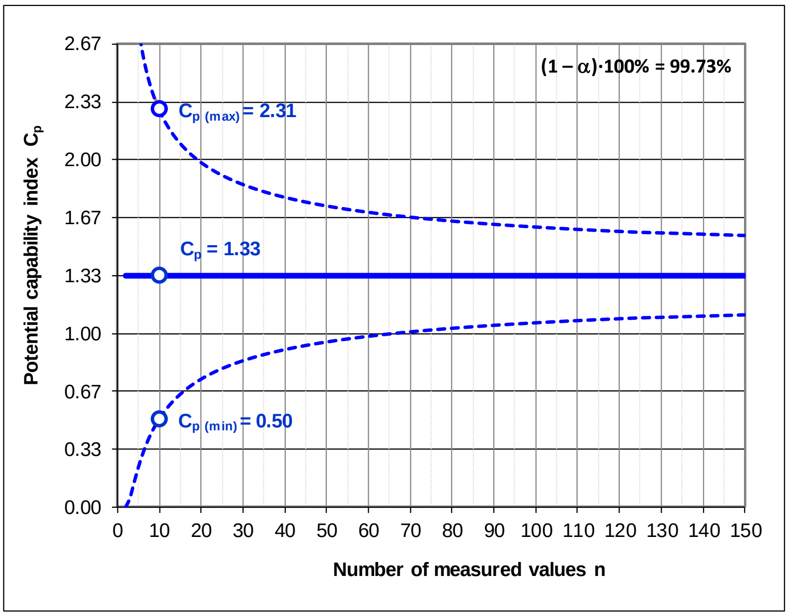


Figure 42: Confidence limits for Cp depending on the number of measurements

Figure 41 shows, that for example at $n = 10$, the actual value for the calculated value $C_p = 1.33$ with 99.73 % probability lies between $C_p = 0.5$ und $C_p = 2.31$, i. e. the actual value of C_p can be significantly smaller than the required minimum value of $C_p = 1.33$ and can be below the lower 99.73 % confidence limit $C_p = 1.08$ which is valid for $n_0 = 125$ measurements.

The concept of ensuring sufficient capability even in the case of low numbers of measurements is based on the premise, that the limit value, which is determined by the lower confidence limit at values, also with $n < 125$ available measurement values does not fall below, meaning, that the lower confidence limit is fixed. This has the consequence, that due to the decreasing number of readings, the width of the confidence interval of the minimum value of C_p is raised.

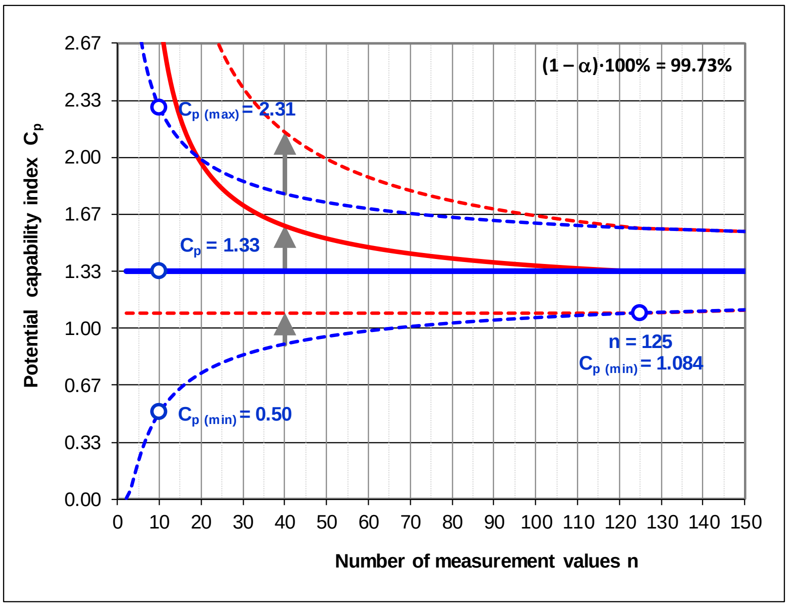


Figure 43: Raising C_p with associated confidence limits depending on the number of readings

The modified values are calculated in accordance with

$$C^*_{p\,(min)} = C_p \cdot \sqrt{\frac{\chi^2_{n_0-1;\,\alpha/2}}{n_0-1}} \qquad C^*_p = C_p \cdot \sqrt{\frac{\chi^2_{n_0-1;\,\alpha/2}}{n_0-1} \cdot \frac{n-1}{\chi^2_{n-1;\,\alpha/2}}} \qquad C^*_{p\,(max)} = C_p \cdot \sqrt{\frac{\chi^2_{n_0-1;\,\alpha/2}}{n_0-1} \cdot \frac{\chi^2_{n-1;\,1-\alpha/2}}{\chi^2_{n-1;\,\alpha/2}}}$$

These calculation rules apply exactly for the potential capability index C_p. Corresponding calculation specifications for C_{pk} can be determined only approximately. However, the numerical differences are of little significance, so that for C_{pk} usually the same calculation rules are used. Accordingly, the following adjustments for C_p and C_{pk} result:

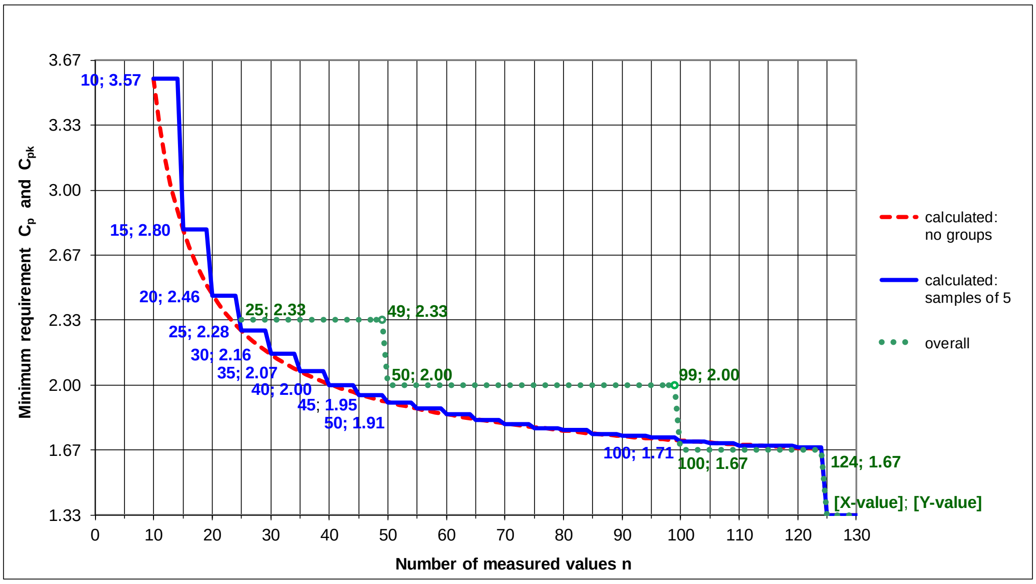


Figure 44: Figure: Minimum values for C_p and C_{pk} at less than 125 available measurement values

The same procedure with the appropriate parameters for the machine capability delivers for C_m and C_{mk} the following adjustments:

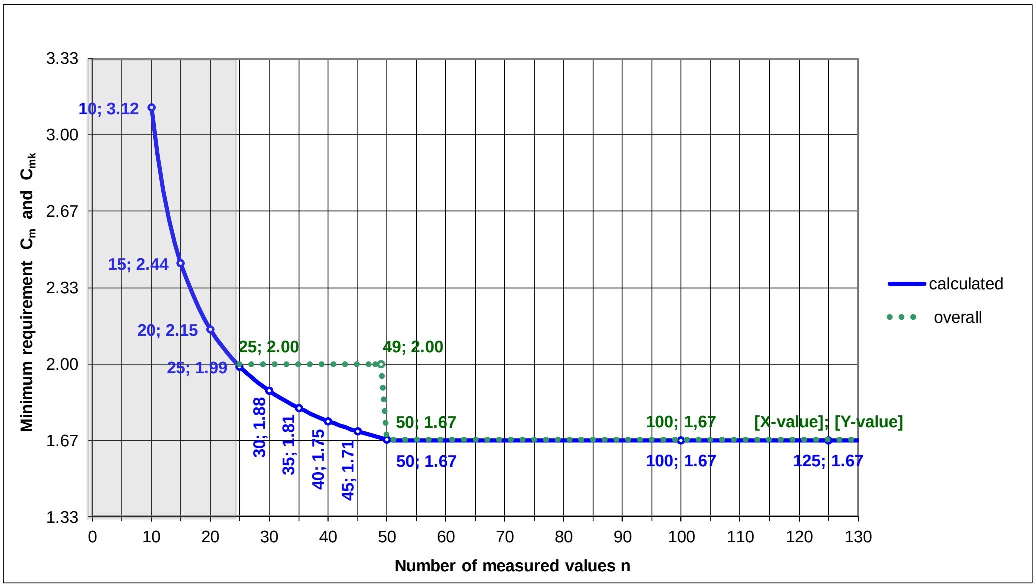


Figure 45: Minimum values for C_m and C_{mk} at less than 50 available measurement values

> *NOTE 1: With certain numbers of available measurements for Cm(k) and Cp(k), in order to attain quantity values which are as closely as possible integer multiples of 0.33, slightly varying confidence levels are occasionally used instead of the 99.73% confidence level. Example Q-DAS software under evaluation strategy BOSCH 2012: Confidence level 99.67% bilateral (corresponding to 99.83% unilateral), so that Cm(k) = 2.0 at n = 25 (minimum number of value measures) and Cp(k) = 2.0 at n = 40 (warning limit for too few readings).*

Numbers under $n = 25$ measurements should absolutely be avoided, because statistical statements for both C_m and C_{mk} and also C_p and C_{pk} are increasingly questionable.

If in exceptional cases, neither a suitable statistical software nor the tables below for common scenarios are available, flat increases are possible:

- Cm(k) will be raised in by a flat 0.33 as soon as the number of available measurements falls below the limit $n = 50$.
- Accordingly Cp(k) will be raised in general by 0.33 in each case, as soon as the number of available measurement falls below the limit $n = 125$, $n = 100$ and $n = 50$ (see Figure).

These general increases usually lead to higher capability requirements as increases which were calculated individually for the actual available number of readings.

> *NOTE 2: If the calculated capability index falls into the area between individual and general determined limit, only that value is valid for the classification "capable" or "not capable" that is actually is actually available for the evaluation and will be used accordingly (i. e. either the individually calculated limit or the limit value taken from the following tables or the generally determined limit).*

All explanations and statements of this chapter regarding process capability are equally applicable to the process performance $P_{p(k)}$ and the short variants $C_{p(k)\text{-}ST}$ and $P_{p(k)\text{-}ST}$ of process capability or process performance.

> *NOTE 3: The minimum requirement for Cp(k)-ST and Pp(k)-ST is already 1.67 at n > 125 values. Therefore in this case, the general increase of 0.33 in the range of 100 < n <125 values is omitted.*

5 measurements per sample		
Number of complete samples	Number of measure-ments in all samples	C_p, C_{pk}, P_p, P_{pk}, $C_{p\text{-}ST}$, $C_{pk\text{-}ST}$, $P_{p\text{-}ST}$, $P_{pk\text{-}ST}$
1	*5*	*7.92*
2	*10*	*3.57*
3	*15*	*2.80*
4	*20*	*2.46*
5	25	2.28
6	30	2.16
7	35	2.07
8	40	2.00
9	45	1.95
10	50	1.91
11	55	1.88
12	60	1.85
13	65	1.82
14	70	1.80
15	75	1.78
16	80	1.77
17	85	1.75
18	90	1.74
19	95	1.73
20	100	1.71
21	105	1.70
22	110	1.69
23	115	1.69
24	120	1.68
≥ 25	**≥ 125**	**1.33**

NOTE: For the short-term indices $C_{p\text{-}ST}$, $C_{pk\text{-}ST}$, $P_{p\text{-}ST}$ and $P_{pk\text{-}ST}$ the minimum requirement 1.67 also applies at more than 125 values.

3 measurements per sample		
Number of complete samples	Number of measure-ments in all samples	C_p, C_{pk}, P_p, P_{pk}, $C_{p\text{-}ST}$, $C_{pk\text{-}ST}$, $P_{p\text{-}ST}$, $P_{pk\text{-}ST}$
1	*3*	*33.10*
2	*6*	*5.97*
3	*9*	*3.88*
4	*12*	*3.16*
5	*15*	*2.80*
6	*18*	*2.57*
7	*21*	*2.42*
8	*24*	*2.31*
9	27	2.22
10	30	2.16
11	33	2.10
12	36	2.06
13	39	2.02
14	42	1.98
15	45	1.95
16	48	1.93
17	51	1.91
18	54	1.88
19	57	1.87
20	60	1.85
21	63	1.83
22	66	1.82
23	69	1.81
24	72	1.79
25	75	1.78
26	78	1.77
27	81	1.76
28	84	1.75
29	87	1.75
30	90	1.74
31	93	1.73
32	96	1.72
33	99	1.72
34	102	1.71
35	105	1.70
36	108	1.70
37	111	1.69
38	114	1.69
39	117	1.68
40	120	1.68
41	123	1.67
≥ 42	**≥ 126**	**1.33**

Machine capability	
Number of measure-ments	C_m, C_{mk}
2	*559.58*
3	*28.90*
4	*11.09*
5	*6.91*
6	*5.21*
7	*4.31*
8	*3.76*
9	*3.39*
10	*3.12*
11	*2.92*
12	*2.76*
13	*2.63*
14	*2.53*
15	*2.44*
16	*2.37*
17	*2.30*
18	*2.25*
19	*2.20*
20	*2.15*
21	*2.11*
22	*2.08*
23	*2.05*
24	*2.02*
25	1.99
26	1.96
27	1.94
28	1.92
29	1.90
30	1.88
31	1.87
32	1.85
33	1.83
34	1.82
35	1.81
36	1.79
37	1.78
38	1.77
39	1.76
40	1.75
41	1.74
42	1.73
43	1.72
44	1.71
45	1.71
46	1.70
47	1.69
48	1.68
49	1.68
≥ 50	**1.67**

Table 6: Requirements for process capability/performance and machine capab. with smaller quantities

(Confidence level 99.67 % bilateral / 99.83 % unilateral below, 5 and 3 readings per sample)

NOTE: Values that do not form a complete dample are not included in the evaluation.

I.2 Grouping of parts and characteristics

To group, similar parts can in particular be suitable for processing together in one operation on the same production unit, so no set-up is required (such as tool change, change of clamping tool), and if this operation has the same and/or similar characteristics (e. g. holes) with the same and/or different degree (e. g. with different diameters) and generates various parts and/or different positions of the same part.

I.2.1 Grouping of different parts: Case B

The existing measurement values of the same or similar characteristics can be grouped together for consecutively manufactured parts with different part numbers, where these characteristics are subject to the same effect, for example, same engine, the same tool (for example, drills, cutters).

When there is diversity of characteristic tolerances, a standardization to a suitable reference value is necessary, such as for example, the respective tolerance T. The tolerances in this case should at least be of the same magnitude.

Standardization for bilateral limited characteristics: $\%x = \frac{x-0{,}5\cdot(LL+UL)}{T} \cdot 100\ \%$

The quantity $\%x$ is the deviation of the current measurement x from the mean value $0{,}5 \cdot (LL + UL)$ of the tolerance interval with the limits LL and UL as a percentage of the tolerance $T = (UL - LL)$. $\%x$ has the nominal value 0% and the limits $\%LL = -50\%$ and $\%LL = +50\%$.

Standardization for zero-limited characterization: $\%x = \frac{x}{T^*} \cdot 100\ \%$

The quantity $\%x$ shows the deviation of the current measurement x of 0 as a percentage of the pseudo-tolerance $T^* = UL - 0 = UL$. $\%x$ has a natural lower limit $\%LL = 0\%$ than the target value with the upper limit $\%UL = +100\%$.

The summary data are calculated using the standardized variables $\%x$, $\%LL$ and $\%UL$ (instead of the non-standardized quantities x, LL and UL) according to Chapter 8 and appropriate capability and performance parameters determined.

NOTE 1: It would be clearer to use indexed notation $\%x_{ik}$, $\%LL_k$, $\%UL_k$ for the standardized quantities and x_{ik}, LL_k, UL_k for the non-standardized quantities as well as $T_k = UL_k - LL_k$ and $T_k^* = UL_k$, where x_{ik} indicates the measurement no. i, which is part of the specification limits no. k, i. e. the limits LL_k and UL_k.

NOTE 2: The results for capability or performance characteristics are independent of whether they are calculated from the standardized or non-standardized quantities.

I.2.2 Grouping of different parts and characteristics of each part: Case C

If despite the summarization of the parts, the number of consecutively manufactured parts remains $n' < 25$, it can be checked whether additional measurements of similar characteristics of a part (for example, several diameters) can be summarized and thus a sufficient data basis can be produced. In general, a suitable standardization is also necessary in this case (see case B, Appendix I.2.1).

The summarized data are analyzed using the normed quantities $\%x$, $\%LL$ and $\%UL$ (instead of unnormed quantities x, LL and UL) according to Chapter 8, and the corresponding capability or performance parameter determined.

I.2.3 Evaluation of the groupability of parts and characteristics

The grouping (summarizing) of parts and characteristics is only useful if the resulting data pool to be statistically analyzed behaves at least approximately as if it came from a single population, although not technically true.

There are no established methods for analysis available yet. Numerical approaches (for example, based on the so-called cluster analysis) are still in the development phase.

Based on suitable processed data, possibly existing structures can nevertheless be identified in the dataset. Such structures can be caused by

- actual process changes
- or data in the data pool, which are unsuitable for the grouping with the other data.

This can only be decided after an analysis of the causes of the identified "irregularities". Therefore, such an analysis is a prerequisite for the proper evaluation of a manufacturing process, that produces, for example, a number of variants, totaling a high number of pieces, but each variant only in small quantities.

To avoid incorrect evaluations, is it crucial for the practical implementation, to identify all possible **relevant** factors and record as additional data (metadata) to each measurement. On the other hand all non-relevant variables should be rejected and not recorded as additional data in order to avoid unnecessarily large and no longer manageable datasets.

EXAMPLE: If there are several variants of a product part, bores with specific diameter variants are produced on a production line. The prepared data show jumps at certain places. The analysis showed that these jumps are correlated with the material of the product part, meaning that they only occur when the material changes from steel to injection molding and vice versa. This would not have been identifiable if only such additional information would have been recorded about the measurements, from which this material transformation does not originate (for example, ambient temperature).

If numerous variants of a particular product are manufactured on one production line and each variant in small numbers and irregular intervals, in particular during the process start-up phase, there is the problem of evaluating the quality capability of the production process.

The following examples illustrate possible approaches for assessing grouping.

Example 1: Stratification (data separation)

Two single components are mechanically linked with adhesive, fixed and sealed. The total weight for both components is measured before and after the application of the adhesive. The difference in weight serves as a characteristic for process control.

Due to large number of variants, the individual variants are produced only in small numbers. Therefore, the measurement results of 2 modules have to be combined, which are manufactured using 2 different types of adhesive on 4 production lines. The specification limits (minimum and maximum weight) are determined exclusively by the type of adhesive used. Considered over an evaluation period of 30 calendar days there are therefore *an average* of 240 readings available. Nevertheless, these are per type of adhesive, module and production line *an average* of only 15 parts per month, which is not sufficient for individual statistical evaluations.

The standardized readings represented in chronological order as a time series graph over seven calendar days of the early start-up phase process (Figure 45), prove to be true on one day as borderline, but are otherwise well centered within ±37.5 % of the respective characteristic tolerances, i. e. the tolerance utilization is less than 75%. In this respect the process seems to largely meet the acceptance criterion

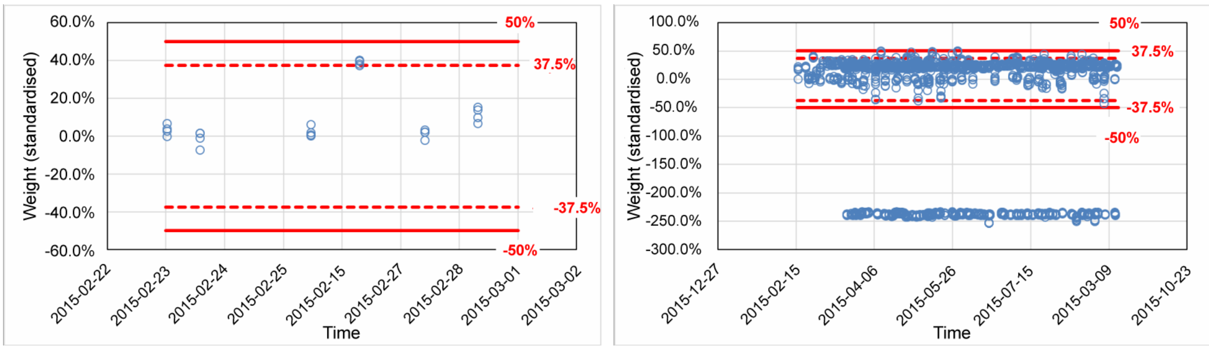


Figure 46: Time series charts “adhesive weight” over 7 calendar days and 7 calendar months

In the time series graph of the standardized values for 7 months, however, a clear structure is discernible: Numerous values lie in a limited range below the lower specification limit, all other values show a clear tendency to the upper specification limit. However, no correlation with the metadata (adhesive type, module, production line) is possible.

Correlations of measurement results with certain characteristic properties and metadata can often be determined by the stratification of the measurements with respect to known parameters such as nominal value, tolerance, material, part number, production line. The measurements are shown separately for these parameters.

In the present case, the stratification according to type of adhesive, shows that the area with readings outside tolerance and the tendency of the measurements towards the upper specification limit is only observed in conjunction with adhesive type 2. But as these two effects occur together, it can be assumed that the adhesive type is not the cause.

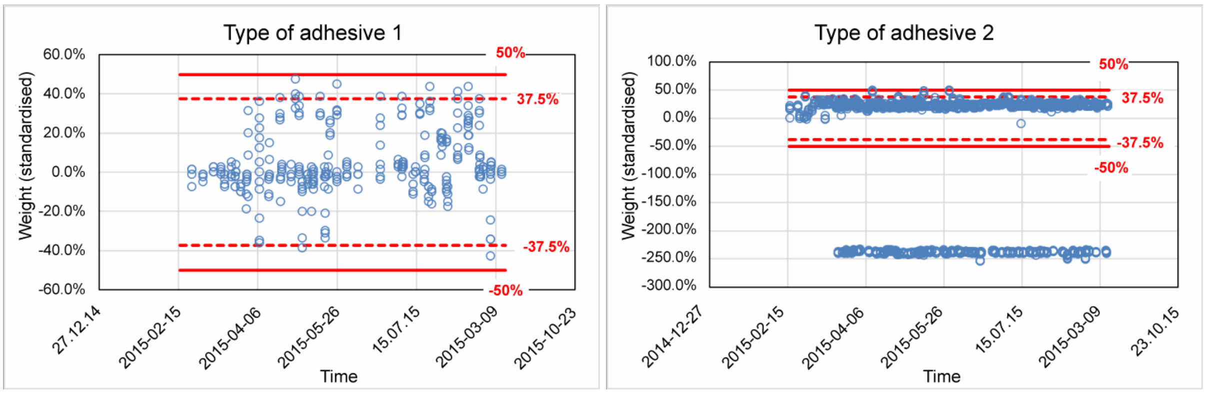


Figure 47: Time series charts “adhesive weight” separated according to adhesive type

The separation according to all available metadata (adhesive type, module, production line) clearly shows that the values outside tolerance are due solely to the production line 3 (Figure 47). It therefore does not make sense to include this line in the group without checking the process location and tolerance and adjusting if necessary.

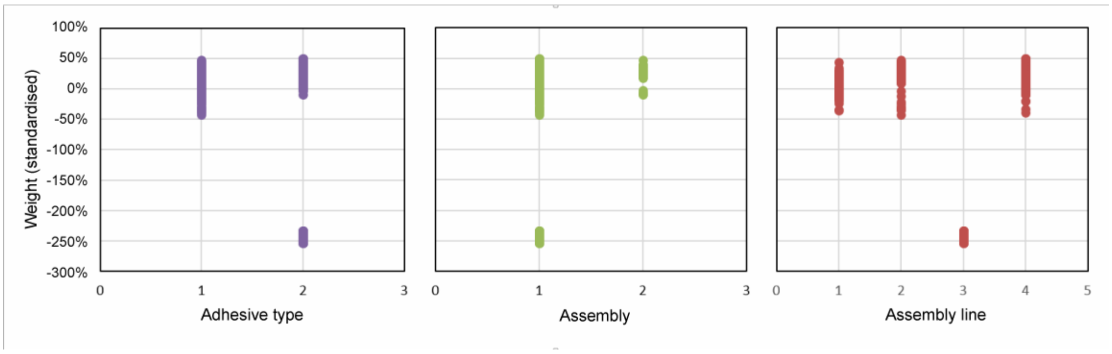


Figure 48: Adhesive weight stratified by type of adhesive, module and production line

The following charts result without production line 3:

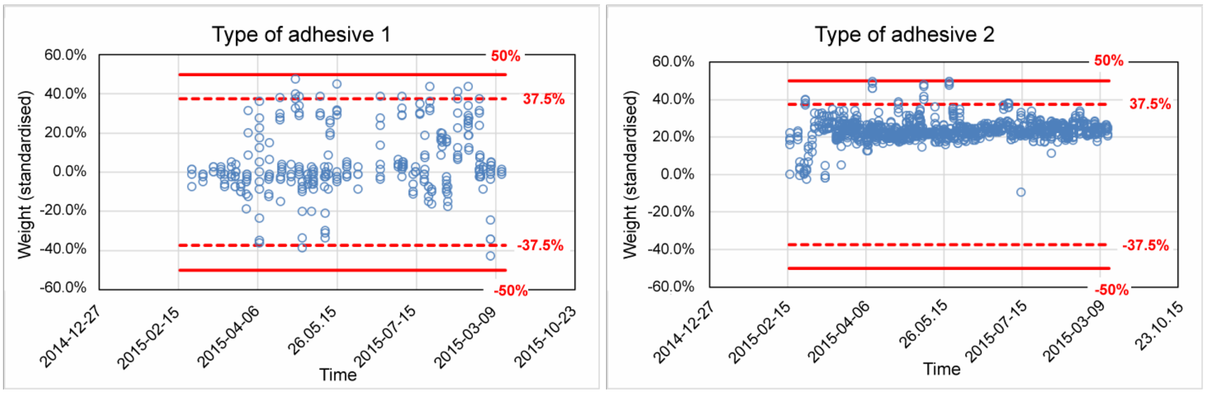


Figure 49: Time series charts "adhesive weight" (production line 3 hidden)

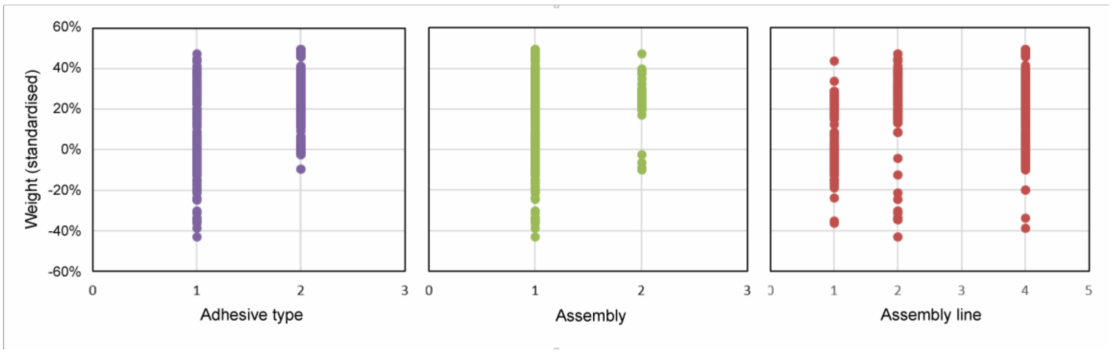


Figure 50: Adhesive weight stratified (production line 3 hidden)

The stratified data show, that the production lines 1, 2 and 4 provide results with a similar process location and variation, so that there is no objection to grouping in a group. Also with regard to adhesive type and assemblies are similar in the deviations, so that at least there are no serious objections to the grouping. Recommended would be to investigate whether a reduction of the variation with adhesive Type 1 reduces the variation and if with adhesive type 2 the process location can be centered better again, which has shifted significantly compared with the initially good centering during the first 30 days of production.

> *NOTE: It is expressly noted that this example merely illustrates the basic procedure. In practice, further analysis and, if appropriate measures are usually required.*

Example 2: Use of mean values and standard deviations

Mean and standard deviation are parameters that can always be determined as a rule, when several measurements are available. They are therefore often used to carry out at least a first, rough estimate.

In the following practical example, for the characteristic "pull-off strength of bonding wire", there are a total $n > 8{,}000$ readings of product parts with $m = 13$ different item numbers. The individual part numbers are represented with a very different number of measurements in the dataset.

The two parameters are determined separately by part numbers (metadata) and the respective standard deviations (y-values) are plotted over the respective mean values (x-values). Frequently such a graph of "abnormalities" already makes it clear, so that no further calculations are required.

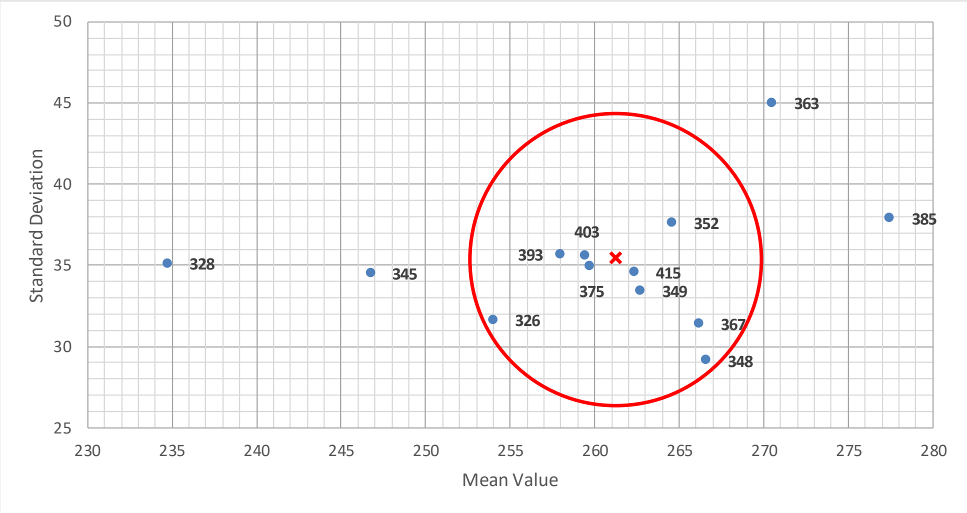


Figure 51: Visualization of grouping based on mean values and standard deviations

Each data point is indicated by the associated (in this example, three-digit) part number.

One way to proceed quantitatively, is to calculate the total mean $\bar{x}$ and the total standard deviation s of all measurements. The data point $(\bar{x}; s)$ represents the middle or center of gravity of all data points $(\bar{x}_i; s_i)$. The radial distance

$$r_i = \sqrt{(\bar{x}_i - \bar{x})^2 + (s_i - s)^2}$$

of a data point i to this center can be used as a quantitative measure for its "abnormality". In this case, the greatest distances result for the data points of the code numbers 328, 345, 363 and 385. This result was already to be expected on a purely visual basis using the diagram.

NOTE: The basic idea of this approach corresponds to a cluster analysis, however, in extremely simplified form. Decision basis for belonging to the cluster (maximum distance for the inclusion of data points, that is, the radius of the circle around the center of the data points) was in this case, also not statistics but plausibility.

The next step should determine the technical causes of these "abnormalities", their effects on the overall result assessed and decided if parts with corresponding part numbers are suitable for grouping or not.

I.3 %T approach: Case D

With less than $n = 25$ measurements, the quality capability can be assessed on the degree of tolerance utilization %T. However, this is appropriate and permissible only under the following conditions:

- 100 % inspection is not possible in principle (destructive test, high unit costs, extremely long measurement times);
- Process visualization until enough parts are available to conduct at least one statistical analysis by Table5 (case A to C) or capabilities according to Chapter 8.

The deviations %x related to the respective tolerance T of the individual measurements from the respective target value (for calculation see Appendix I.2.1) are **not** used in this case to calculate statistical parameters such as capability and performance indices. Instead, the %x-individual values will be assessed on the basis of the criterion of whether in the presence of negligible uncertainty U all

%x-values lie

- at bilateral limited characteristics in the range -37.5 % to +37.5 %
- and at zero-limited characteristics in the range 0 % to 75 %

so that max. 75 % of the tolerance interval T are used. This 12.5% - or 25% -distance to the specification limit reduces the risk to a certain degree that the process produces outside of tolerance in the time interval between two tests. However, it is expressly understood that this is merely a plausible, but in no way constitutes a statistically reliable risk reduction: The %T approach is purely **a visualization method**, which only provides a snapshot of the current situation and is not necessarily indicative of future process behavior and no prognosis on a statistical basis.

NOTE 1: The tolerance utilization may be limited to less than 75%, e. g. 60%. In this case, the %x-individual values must be in the range -30% to + 30% (bilateral limited) or 0 to 60% (zero-limited). This contributes i. a. to further risk reduction.

NOTE 2: Taking into account the uncertainty U of the test modality can be omitted if this is less than 10% of the smallest tolerance to be tested with this test modality.

NOTE 3: Sample taking and interval must be reasonable for the observed (or expected) process behavior, the significance of the observed characteristics and the associated testing expenditure.

NOTE 4: It is not allowed, to attain or establish an evaluation only by the %T approach that despite adequate batch size, only a few parts are measured.

NOTE 5: Statements regarding quality capability of a process based on the relative tolerance utilization %T are basically not comparable to "classic" capability statements based on statistical capability indices. Comparative statements about capability parameters are only meaningful and permitted if they have been determined by the same method.

NOTE 6: The acceptance criteria %x ≤ 75% appears to be plausible with a view to the limit of 1.33 of statistical parameters such as Cp and Cpk since this level is also defined for more than 75% tolerance utilization. In diametric contrast to the ***non-statistical*** *%T approach, here, however, it is statistically supported through a distribution model that 99.73% of the process results are expected to be within the 75% -tolerance interval. This is not ensured at the %T approach in any way.*

J Capability indices with discrete characteristics

The determination of capability parameters such as C_p and C_p according to Chapter 8 assumes a measurable characteristic. Such a characteristic can, in principle, assume any value and is therefore referred to as a continuous characteristic. Measurement results are limited by the measuring range and the resolution of the measurement system.

However, there are processes where this condition is not met. For example, in the printed circuit board mounting or soldering, mistakes can occur which, however can be *counted*, but not in a real sense *be measured*, for example placement errors (incorrect or missing component, incorrect part orientation) or soldering defects (cold solder joint, short circuit, lack of contact). Such a property (for example, solder defects) is called an attributive characteristic, which often has only two levels (e. g. present, not present). The counting of such an attributive characteristic over several parts is referred to as a discrete characteristic (for example, 10 parts incorrectly), whose expression can only be changed in integer steps, i. e. *discrete* (see also [Booklet 2]).

[ISO 22514-5] provides an approach which allows indication of a capability or performance characteristic, even in such cases. The proportion $\hat{p} = \frac{k}{n}$ of the number of defective parts k in a sample related to the sample size n is used as an estimator for the probability p of finding defective parts in the sample.

Unlike some older approaches, [ISO 22514-5] takes the confidence interval $p_L \le \hat{p} \le p_U$ this estimator into consideration, where p_L and p_U designate the lower or upper confidence limit at a given confidence level $1 - \alpha$. As error rates cannot be less than 0, only the upper confidence limit p_U is relevant. p_U provides the statistically worst case for $\hat{p}$ and is therefore used instead of $\hat{p}$ for further calculations.

Characteristics with two levels are some of the most common cases which occur in practice, and can generally be described by the binomial distribution. The *exact* confidence limits for the parameter p of the binomial distribution are determined using the beta distribution[18]:

$$p_U = Beta^{-1}(1 - \alpha;\ k + 1;\ n - k)$$

$Beta^{-1}$ refers to the inverse function of the beta distribution 18 and provides the quantile at certain probabilities 1 - α. With this type of calculation, p_U represents the upper limit of the so-called Clopper-Pearson-interval in this type of calculation. .

NOTE 1: [ISO 22514-5] provides in the case $k > 0$, an approximate approach based on the normal distribution calculation of p_U (also known "standard interval" or "forest interval"). The applicability of this approximation is, however, often limited to the area $k \ge 50$ and $n - k \ge 50$ in the literature and may prove problematic even in this area. For the primary practice-relevant range $0 \le k \le 2$, the approximation is therefore not unequivocally suitable. Therefore only the exact approach shown (Clopper-Pearson interval) is applicable here, which is also applicable in the common case $k = 0$ (no faulty part in the sample).

In order to determine a characteristic value, p_U is considered as an estimator for the fractions nonconforming of a normal distribution. Accordingly, $1 - p_U$ is the probability of samples without defective parts. The quantile u_{1-p_U} of the standard normal distribution represents the limit of this error-free area and is used analogously to a unilateral upper specification limit. The distance to the center position $\mu = 0$ of the standard normal distribution results in terms of their half variation $3 \cdot \sigma$ for $\sigma = 1$ according to the standard calculation rules to capability or performance indices

$$\begin{matrix} C_{pk} \\ P_{pk} \end{matrix} = \frac{u_{1-p_u} - \hat{\mu}}{3 \cdot \sigma} = \frac{u_{1-p_u}}{3}$$

[18] Calculation for example with EXCEL spreadsheed function =BETA.INV(1-α; k+1; n-k)

If no control chart exists, an unambiguous evaluation of the process stability is usually not possible. In this case, the determined characteristic value should be considered as process performance and designated by P_{pk}.

Table 7 contains some selected results for p_U and P_{pk} at $k = 0$, $k = 1$ and $k = 2$ defective parts in the sample and the confidence level 95% depending on the sample size n. Figure 51 represents these results graphically.

Sample size n	Confidence limit p_U (k = 0)	Performance index P_{pk}	Confidence limit p_U (k = 1)	Performance index P_{pk}	Confidence limit p_U (k = 2)	Performance index P_{pk}
5	45.1 %	0.04	65.7 %	0.00	81.1 %	0.00
6	39.3 %	0.09	58.2 %	0.00	72.9 %	0.00
7	34.8 %	0.13	52.1 %	0.00	65.9 %	0.00
8	31.2 %	0.16	47.1 %	0.02	60.0 %	0.00
9	28.3 %	0.19	42.9 %	0.06	55.0 %	0.00
10	25.9 %	0.22	39.4 %	0.09	50.7 %	0.00
20	13.9 %	0.36	21.6 %	0.26	28.3 %	0.19
30	9.5 %	0.44	14.9 %	0.35	19.5 %	0.29
40	7.2 %	0.49	11.3 %	0.40	14.9 %	0.35
50	5.8 %	0.52	9.1 %	0.44	12.1 %	0.39
60	4.9 %	0.55	7.7 %	0.48	10.1 %	0.42
70	4.2 %	0.58	6.6 %	0.50	8.7 %	0.45
80	3.7 %	0.60	5.8 %	0.52	7.7 %	0.48
90	3.3 %	0.61	5.2 %	0.54	6.8 %	0.50
100	3.0 %	0.63	4.7 %	0.56	6.2 %	0.51
200	1.5 %	0.72	2.3 %	0.66	3.1 %	0.62
300	1.0 %	0.78	1.6 %	0.72	2.1 %	0.68
400	0.7 %	0.81	1.2 %	0.75	1.6 %	0.72
500	0.6 %	0.84	0.9 %	0.78	1.3 %	0.75
600	0.5 %	0.86	0.8 %	0.80	1.0 %	0.77
700	0.4 %	0.88	0.7 %	0.82	0.9 %	0.79
800	0.4 %	0.89	0.6 %	0.84	0.8 %	0.81
900	0.3 %	0.90	0.5 %	0.85	0.7 %	0.82
1.000	0.3 %	0.92	0.5 %	0.86	0.6 %	0.83
2.000	0.1 %	0.99	0.2 %	0.94	0.3 %	0.91
5.000	0.1 %	1.08	0.1 %	1.04	0.1 %	1.01
10.000	0.030 %	1.14	0.047 %	1.10	0.063 %	1.08
20.000	0.015 %	1.21	0.024 %	1.16	0.031 %	1.14
50.000	0.006 %	1.28	0.009 %	1.24	0.013 %	1.22
100.000	0.003 %	***1.34***	0.005 %	1.30	0.006 %	1.28
200.000	0.001 %	1.39	0.002 %	***1.36***	0.003 %	***1.33***
500.000	0.001 %	1.46	0.001 %	1.43	0.001 %	1.40

Table 7: Confidence limit and performance index, depending on the sample size

(Confidence level 95 %, k = 0, 1, 2 defective parts in the sample)

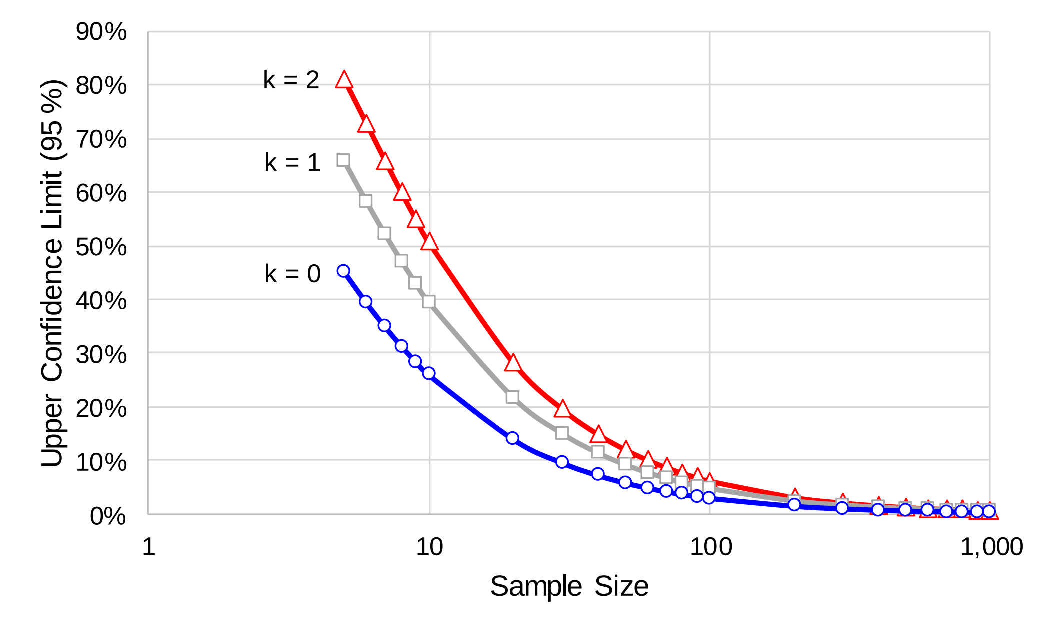


Figure 52: Upper limit of the Clopper-Pearson intervall, depending on the sample size
(Confidence level 95 %, k = 0 … 2 defective parts in the sample)

Assessment of the process

Table 7 shows that sample sizes of the order $n = 100{,}000$ and more would be needed to prove $P_{pk} = 1.33$ confidence level 95 % statistically. This is in practice usually not feasible.

However, based on realistic sample sizes up to the order of at most a few hundred parts, at best, $P_{pk} = 0.6$ to $P_{pk} = 0.85$ can be reached.

EXAMPLE: Sample of n = 200 parts and k = 0 defective parts

According to Table 7 the process performance Ppk = 0.72 results at confidence level 95 %. This means that the process with 95 % probability ***delivers no more*** *than 1.5 % defective parts. But this also means that the process* ***may possibly*** *return many fewer defective parts, which, however, is not statistically proven.*

If the sample size for example would be increased n = 600 parts and k = 0 defective parts are found, it would be statistically demonstrated, that the process with 95 % probability delivers ***at most*** *0.5 % defective parts (Ppk = 0.86), at k = 1* ***at most*** *0.8 % (Ppk = 0.80) and at k = 2* ***at most*** *1 % (Ppk = 0.77). However, this in turn means also that the process* ***may*** *return much less than 0.5% and 0.8% and 1% defective parts, but this is not statistically proven.*

The example shows that the formal fulfillment of a criterion such as $P_{pk} \geq 1.33$ with discrete characteristics is economically almost impossible and factually does not make a lot of sense. Instead, it is recommended to set a “reasonable” number of defective parts similarly to sampling plans, which the process may deliver undetected at a maximum, and to determine the sample size accordingly. This procedure is not covered by national or international standards or guidelines, however, and must therefore be adapted in each individual case with internal or external customers.

NOTE 2: [ISO 22514-5] indicates the number of defective parts in the case $k > 0$ based on the sample size as a quality level Q_P of the process und uses Q_P like a parameter. When using control charts, an average quality level Q_P can be determined from the values of the control chart. This is not consistent with the approach in the case of $k = 0$, in which the upper confidence limit is used as a quality level Q_P, so that $k = 0$ usually leads to a less favorable quality level Q_P as $k > 0$. Because of this inconsistency, the use of the term “quality level” is not recommended.

K Characteristics limited to only one side

In connection with CDQ 0301 and also with the procedures of this Booklet and Booklet No. 10, the question often arises what is a "unilaterally limited characteristic". The relevant standards on quality management and statistical terms do not currently contain a corresponding definition.

Typically, the characteristics limited on one side are geometric characteristics such as straightness, roundness, flatness, cylindricity, parallelism, perpendicularity, angularity, concentricity, and circular run-out. These are examples of characteristics of form, orientation, location or run-out whose tolerancing is described in ISO 1101. For example, they are defined by the distance from a reference point, a reference line, or a reference surface. Due to the mathematical meaning of the word "absolute value", such amounts of distances or deviations can only assume values greater than or equal to zero. The absolute value is the (non-negative) distance to the number zero.

The point is that the development has defined only a lower limit or only an upper limit which is necessary, for example, due to construction. Then no tolerance T is defined. For the calculation of the critical indices for "Unilaterally limited characteristics" see the corresponding section in 8.1.

The target value does not necessarily have to be zero for a zero-limited characteristic. The roughness depth Rz is a zero limited characteristic; however, due to the desired oil adhesion for sliding surfaces, it may be necessary to specify both a lower limit greater than zero and an upper limit.

Naturally limited characteristics

In the literature, the term "physically conditioned" or "technologically conditioned" is often found in brackets as a supplement to the wording "unilaterally limited characteristic".

However, this is not always correct. For example, the concentration of a chemical solution, i. e. a mixture of substances, can only be between 0 and 1. The reason for this is not physics or chemistry or technology, but the definition of concentration as a quotient (see below). Seen in this way, the word "natural" in the sense of "by definition" or "given by nature" is neutral without referring to a scientific field or a scale.

Due to its definition, a naturally limited characteristic already has a lower limit that a determined value of this characteristic cannot fall below or an upper limit that the value cannot exceed. Both limits can also be given by definition.

Probably hardly anyone thinks of the term "naturally limited characteristic" at first that physical basic quantities such as length, mass, time, weight, force, current, temperature (in Kelvin), amount of substance and luminous intensity or quantities derived from them such as area, volume, density, pressure, torque or speed cannot become smaller than zero. The same applies to the electrical quantities voltage, resistance and frequency.

They often "hide" behind technical designations such as height, width, depth, distance (from objects), diameter (of objects, drill holes), layer thickness (of coatings), roughness (of surfaces), tear strength / tensile strength (of wires or adhesive joints), burst pressure (of membranes), Adhesion (of coatings, lacquers), prevail torque or breakaway torque (of screw connections), dielectric strength (of insulators), duration, running time, time difference, time interval (of electrical, optical, acoustic signals).

Naturally two-sided limited characteristics

The values of characteristics defined as ratios (quotients, fractions) are usually expressed in percent (%), per mille (‰), parts per million (ppm) or parts per billion (ppb). Examples are

- the material proportion Rmr of a surface profile at a defined cutting height
- the concentration or purity (proportions in solids, liquids, gases): quotient of masses, volumes, quantities of substances or numbers of particles
- the efficiency (mechanical or electrical machines/devices)
- the degree of reflection (the reflectivity of surfaces): ratio of reflected and incident intensity

Other reasons for natural limits may be specified test methods and scales, e. g. in the case of hardness (mechanical resistance to penetration of a specimen into a material) or the pH-value.

Symbol directory

$\%GRR$	Total variation width of a measurement process based on the tolerance of the characteristic or on the total variation width of the manufacturing process *(MSA,*
c_4	Factor for the determination of $\hat{\sigma}$ from the mean standard deviation $\bar{s}$ *(in older literature also referred to by a_n)*
C_g, C_{gk}	Potential and critical measurement process capability index *(MSA, Procedure 1)*
C_m, C_{mk}	Potential and critical process capability index
C_p, C_{pk}	Potential and critical process capability index *(long-term)*
C_{p-ST}, C_{pk-ST}	Potential and critical process capability index *(short-term)*
d_2	Factor for the determination of $\hat{\sigma}$ from the mean range $\bar{R}$
i	Number (index) of the measurement within a sample; $1 \leq i \leq n$
j	Number (index) of the measurement within all measurements; $1 \leq j \leq m \cdot n$
k	Number (index) of the sample within all samples; $1 \leq k \leq m$
LCL	Lower control limit *(engl. **L**ower **C**ontrol **L**imit)*
LL	Lower specification limit, minimum value *(**L**ower **L**imit)*
m	Number of samples
n	Number of readings per sample *(sample size)* or in a value set
n′	Number of parts *(if different from n)*
P_m, P_{mk}	Potential and critical machine performance index *(according to ISO 22514-3 instead of C_m and C_{mk})*
P_p, P_{pk}	Potential and critical process performance index *(long-term)*
P_{p-ST}, P_{pk-ST}	Potential and critical process performance index *(short-term)*
p_L	Lower confidence limit
p_U	Upper confidence limit
R	Range of a value set
R_k	Range of sample no. k
$\bar{R}$	Mean of ranges
s	Empirical standard deviation
s_{total}	Standard deviation of all individual values
$s_{\bar{x}}$	Standard evaluation of the mean values of m samples
$\bar{s}$	Mean standard deviation from m samples of the same size
$\overline{s^2}$	Mean variance; mean squared standard deviations
σ	Standard deviation of the population
$\hat{\sigma}$	Estimate value for the standard deviation of the population
T	Tolerance of a characteristic
u_{1-p}	Quantile of the standard normal distribution for the probability $1-p$
UCL	Upper control limit
UL	Upper limit

x_i	Single value no. i of a value set
x_{ik}	Single value no. i in sample no. k
x_{max}	Largest individual value of a value set *(maximum)*
x_{min}	Smallest individual value of a value set *(minimum)*
$\bar{x}$	Arithmetic mean
$\bar{x}_k$	Arithmetic mean of the individual values in sample no. k
$\bar{\bar{x}}$	Mean of mean values (total mean; grand mean)
$\tilde{x}$	Median
$\tilde{x}_k$	Median of the individual values in sample no. k
$\bar{\tilde{x}}$	Mean of medians
$\hat{X}_{0,135\ \%}$	Estimate value for the 0.135 % quantile of the value distribution of the population
$\hat{X}_{50\ \%}$	Estimate value for the 50 % quantile of the value distribution of the population
$\hat{X}_{99,865\ \%}$	Estimate value for the 99.865 % quantile of the value distribution of the population

Other symbols used only in individual chapters or the use of symbols with different meanings are defined in the appropriate context.

	Booklet 9 *(2005)*	DIN 55319 *(withdrawn)*	ISO 21747 *(withdrawn)*	qs-STAT *(as of 06/2016)*	ISO 22514	Booklet 9 *(2016), (2019)*
Lower **L**imit **L**ower **S**pecification **L**imit	LL	L	L	LSL	L	LL
Upper **L**imit **U**pper **S**pecification **L**imit	UL	U	U	USL	U	UL
Lower **C**ontrol **L**imit	LCL	–	–	LCL	–	LCL
Upper **C**ontrol **L**imit	UCL	–	–	UCL	–	UCL
Lower quantile (0.135 %)	$\hat{Q}_{0.00135}$	$\hat{Q}_{0,00135}$	$X_{0.135\ \%}$	$Q_{un\ 3}$	$X_{0.135\ \%}$	$X_{0.135\ \%}$
Upper quantile (99.865 %)	$\hat{Q}_{0.99865}$	$\hat{Q}_{0.99865}$	$X_{99.865\ \%}$	$Q_{ob\ 3}$	$X_{99.865\ \%}$	$X_{99.865\ \%}$

Table 8: Variants of common symbols and abbreviations

Terms and definitions

NOTE 1: The following definitions were taken from the respective cited standards and guidelines. Related comments were taken only in exceptional cases, if they were directly relevant and/or indispensable for understanding a concept. Otherwise it is referenced in respect to the comments and examples to the respective standard or directive.

*NOTE 2: Editorial remarks are **not** part of the respective standard or guideline.*

NOTE 3: Term definitions are mainly taken from [ISO 22514-1], [ISO 3534-2], [ISO 3534-1], [ISO 9000] and [VIM]. In some cases, the same term is listed with several definitions from various standards and directives, provided that the definitions do not appear entirely consistent.

NOTE 4: Terms whose definitions are included in the compilation are displayed in bold when used in definitions of other terms.

Capability

Suitability of a object *(for example **product**, service, **process**, person, organization, system, resource)* for realization of a result, which will fullfill the **requirements** of this result *(with reference to [ISO 9000, 3.6.12])*

Capability

Suitability of an organization, a system or an **process** for realisation of a **product**, which will fullfill the **requirements** of this product [ISO 22514-1. 3.3.2]

Capability index *(English capability index):* see **process capability index**

Characteristic

Distinguishing property

NOTE 1: A characteristic can be inherent or assigned.

NOTE 2: A characteristic can be of qualitative or quantitative nature.

NOTE 3: There are different classes of characteristics, e. g.:

- *physical, for example mechanical, electrical, chemical or biological characteristics;*
- *sensory, e. g. regarding smell, touch, taste, sight, hearing;*
- *behavioral, e. g. decency, honesty, truthfulness;*
- *time-related, for example punctuality, reliability, availability;*
- *ergonomic, for example physiological or characteristics related to human safety;*
- *functional, for example top speed of an airplane.*

[ISO 3534-2, 1.1.1]

Confidence interval

Range estimator (T_0, T_1) for a parameter θ, wherein the **parameters** T_0 and T_1 are interval limits and it applies, that $P[T_0 < \theta < T_1] \geq 1 - \alpha$.

NOTE 2: Connected to this confidence interval is the associated parameter $100 \cdot (1 - \alpha)$ %, in which α in general is a small number. This parameter, which is called the confidence coefficient or confidence level often has the value 95 % or 99 %. The inequation $P[T_0 < \theta < T_1] \geq 1 - \alpha$ is true for a well specified, but unknown value θ of the population.

[ISO 3534-1, 1.28]

EDITORIAL NOTE: P denotes a probability

Confidence level see **confidence interval**, Note 2

Conformity

Fulfilling a **requirement** [ISO 9000, 3.6.11]

Conformity evaluation

Systematic **examination** about the degree to which an **entity** fulfills special **requirements** [ISO 3534-2, 4.1.1]

Continuous characteristic

Characteristic whose characteristic values are the **measuring values** of a physical **quantity** (for example, weight, length, current, temperature)

EDITORIAL NOTE: Often imprecisely called "variable characteristic"; see also "continuous scale"

[ISO 3534-2, 1.1.4]

Determination

Activity to determine one or more **characteristics** and its characteristic values [ISO 9000, 3.11.1]

Discrete characteristic

Characteristicwhose characteristic values are counted measurands in a countable **unit** (e. g. good/bad, right/wrong, red/green/blue)

EDITORIAL NOTE: Often imprecisely called "attributive characteristic"; see also "discrete scale"

[ISO 3534-2, 1.1.5].

Process characteristic in control *(English translation of DIN 55350-11 not available)*

Process characteristic, in which the parameters of the distribution of the characteristic values practically do not change or only change in a known manner or in known limits [DIN 55350-11, 3.11.1]

Empirical

Derived from experience, from observations, determined experimentally, without the use of a mathematical model

EDITORIAL NOTE: In statistics, the adjective "empirical" is used to distinguish variables determined from measurement data, such as mean value $\bar{x}$ and standard deviation s, from the corresponding variables μ and σ of a probability distribution. The different spelling of these quantities with Latin and Greek letters serves the same purpose.

Entity

That which can be individually described and considered [ISO 3534-2, 1.2.11]

EDITORIAL NOTE: Not to be confused with "unit" (see [VIM, 1.9])

Estimate

Observed value of an **estimator** [ISO 3534-1, 1.31]

Estimation

A method that is gaining a statistical representation of a **population** from drawing a **random sample** from this population..

NOTE 1: In particular, the method justifies the estimation that leads from an estimator to a special estimate.

[ISO 3534-1, 1.36]

Estimator

Statistic which is used to **estimate** a parameter Θ [ISO 3534-1, 1.12]

Indication

From a **measuring device** or **measuring system** delivered **quantity value** [VIM, 4.1]

Influence quantity

Quantities, engaged in a direct **measurement** which does not affect the quantity which is being measured, but affects the relationship between the **indication** and the **measurement result**

[VIM, 2.52]

Inspection

Conformity assessment through observation and assessment, accompanied — if applicable — by **measurement**, testing or comparison [ISO 3534-2, 4.1.2]

Lower specification limit

Limit indicating the lower limiting value [ISO 3534-2, 3.1.5]

Measurand

Quantity which is to be measured [VIM, 2.3]

Measurement

Process in which one or more **quantity values**, that can be reasonably assigned to a **quantity**, are experimentally determined

NOTE 2: A measurement means comparing sizes and includes counting

[VIM, 2.1]

Measuring instrument

Device, which is used alone or in conjunction with additional facilities for the performance of **measurements** [VIM, 3.1]

Measurement process

Set of activities for **determining** a **quantity value** [ISO 9000, 3.11.5]

Measurement result

Set of **quantity values** that are assigned to a **measurand** along with any available relevant information [VIM, 2.9]

Measurement uncertainty

Not negative parameter that characterizes the variation **of values** that is attached to the **measurand** on the basis of information used [VIM, 2.26]

Measurement value

Quantity value which represents a **measurement** [VIM, 2.10]

Measuring system

Combination of **measuring devices** and often other devices and, if necessary reagents and utilities, which are arranged and adapted to provide information to **obtain readings** within certain intervals for **quantities** of certain **types**

> *NOTE: A measuring system can consist of a single measuring device*

[VIM, 3.2]

Observed value

Value obtained of a property, which is connected with an element of a **sample**

[ISO 3534-1, 1.4]

<u>Order statistic</u>

Statistic determined by its ranking in a nondecreasing arrangement of random variables

[ISO 3534-1]

Parameter

Index of a family of distributions

> *NOTE 1 The parameter may be one-dimensional or multi-dimensional.*
>
> *NOTE 2 Parameters are sometimes referred to as location parameters, particularly if the parameter corresponds directly to the mean of the family of distributions. Some parameters are described as scale parameters, particularly if they are exactly or proportional to the standard deviation of the distribution. Parameters that are neither location nor scale parameters are generally referred to as shape parameters.*

[ISO 3534-1]

> *EDITORIAL NOTE: Parameters of location (expected value, median), parameters of variability (variance, standard deviation, coefficient of variation) and shape parameters (skewness, kurtosis, excess) are also called functional parameters.*

Performance index

Parameter that indicates the **performance measure** with respect to stipulated **specifications**

[ISO 22514-1, 3.2.3]

Performance measure

Statistical **measurand** of the result for a **characteristic** from a **process** from which it does *not have to be* shown, that it is a **stable process**

[ISO 22514-1, 3.2.2]

<u>Place value</u>

Rounding place: The position of a number symbol of the decimal system where the last digit should be after the rounding. Note: In case of measurement and calculation results, no zeros should appear after the rounding point when numbers are rounded. The value of the rounding place is called **place value**.

Analogous translation from [DIN 1333]; see also [ISO 80000-1].

Population

Population of the considered **entities** [ISO 3534-2, 1.2.1]

Population parameter

Summary measure of the values of some characteristic of a population

> *EXAMPLE Population mean = μ; population standard deviation = σ*
>
> *NOTE Population parameters are usually symbolized by lower case Greek letters in italics.*

[ISO 3534-2]

Process

Set of associated or mutually influencing activities, which uses entries to achieve an intended result [ISO 9000, 3.4.1]

Process capability index

Parameter which indicates the **capability** with respect to given **specifications** [ISO 22514-1, 3.3.6]

Process characteristic

Inherent **characteristic** of a **process** [ISO 22514-1, 3.1.8]

> *EDITORIAL NOTE: Process characteristics are necessary characteristics to ensure the **conformity** of the **product characteristics**. These are characteristics that are transmitted to facilities and equipment, and not on the **product**.*
>
> *For the term "inherent", see also the editorial note with regard to the term "product characteristic".*

Product

Result of a **process** [ISO 22514-1, 3.1.4], [ISO 3534-2, 1.2.32]

Product characteristic

Inherent **characteristic** of a **product** [ISO 22514-1, 3.1.7]

> *EDITORIAL NOTE: "Inherent" means "an inherent unit" (for example, physical properties such as weight, size, power consumption of a product); therefore an inherent characteristic may be a **quality characteristic**, but not a "mapped" characteristic (such as, e. g., price, the owner).*

Product characteristic in controlN2)

Product characteristic parameters of the distribution of the characteristic values, which virtually do not change or only change in a known manner or within known limits.

N2) National footnote: ISO 21747: 2006, 3.1.1.6 used the English terms "stable process" and "process in a state of statistical control" interchangeably, which DIN ISO 21747 translated as "stable process" and "dominated process".

Deviating from that, ISO 22514-1: 2014 3.1.21 designated only the behavior after ISO 21747:2006, 3.1.1.6, notes 1 to 3, as "stable process" and "process in a state of statistical control", which DIN ISO 22514-1 translated as "stable process".

The behavior according to ISO 21747: 2006, 3.1.1.6, note 4, is designated as ISO 22514-1: 2014 3.1.20, however, as a "product characteristic in control", which is translated as "dominated product characteristic".

This important change is not yet considered in DIN ISO 3534-:2013-12. 2.2.7.

[ISO 22514-1, 3.1.20]

Quality capability

Suitability of an organization or parts of an organization (e. g. people, procedures, processes, equipment) for realizing a result that will meet the quality requirements of this result *(with reference to [ISO 22514-1, 3.3.2] and [ISO 9000, 3.6.12])*.

Quality characteristic

Inherent **characteristic** of a **product**, a **process** or system related to a **requirement** [ISO 22514-1, 3.1.9]

EDITORIAL NOTE: For the term "inherent", see also the editorial note for the term "product characteristic".

Quantity

Property of a phenomenon, a body or a substance wherein the property has a value which can be expressed by a number and a reference [VIM, 1.1]

Quantity type

Aspect which is common with comparable **quantities** [VIM, 1.2]

Quantity value

Numerical value and reference, which together specify a **quantity** quantitatively [VIM, 1.19]

Random cause

Cause of the variation which is constantly inherent in a **process** [ISO 22514-1, 3.1.19]

Random sample

Sample, which has been selected randomly [ISO 3534-1, 1.6]

Requirements

Requirement or expectation that or which is stipulated, commonly provided or mandatory
[ISO 9000, 3.6.4]

Resolution

Smallest change of a **measurand**, that causes in the corresponding **display** a noticeable change [VIM 4.14]

Resulting process distribution

Time-dependent distribution model that reflects the instantaneous distribution of the characteristic under consideration, and the changes of its location, dispersion and shape parameters during the time interval of process observation. According to [ISO 22514-2].

Sample

Subset of a **population** which consists of one or more **selection units**. [ISO 3534-1, 1.3]

Sampling unit

One of the individual parts, of which a **population** is composed [ISO 3534-1, 1.2]

Specification

Document that specifies **requirements**

> *NOTE: A specification may refer to activities (for example process document, process specification and test specification), or products (for example, product specification, performance specification and drawing).*

[ISO 9000, 3.8.7]

Specification interval

Area between the **limits maximum value** and **minimum value** [ISO 22514-1, 3.1.14]

> *EDITORIAL NOTE: The limits are also referred to as specification limits.*

Specification limit

For a **characteristic** of a fixed limiting value [ISO 3534-2, 3.1.3]

Stable process *(English translation of DIN 55350-11 not available)*

Process whose essential characteristics are **stable process characteristics** [DIN 55350-11, 3.11.2]

Stable process *(process in a state of statistical control)*

Process which is only **subject to random scatter causes**

> *NOTE 1: A stable process behaves in general as if **the samples** drawn from the process are at any time simple **random samples** n from the same **population**.*
>
> *NOTE 4: In some processes, the mean value of a **characteristic** may drift, or the standard deviation increase, for example due to tool wear or due to a decreasing concentration of a solution. A progressive change of the mean value or the standard deviation of such a process is considered to be the result of systematic causes and not as a result of **random variation causes**. Thus no simple random samples are obtained from the same population.*

[ISO 3534-2, 2.2.7]

> *EDITORIAL NOTE: The English original version of this term is identically defined in ISO 3534-2 (2006) and ISO 21747 (2006); DIN ISO 3534-2 contains the newer German translation (2013); DIN ISO 21747 contains the older German translation (2007) and uses the terms “**stable process**” and “**dominated process**” synonymously.*

Stable process; process in a state of statistical control

Process which (with regard to its variation) is only subject to **random causes**

NOTE 1: A stable process will behave in general as if the samples are random samples at any time with simple sampling from the same population.

NOTE 4: In some processes, the expected value of the characteristic can change, or the standard deviation can be increased. The reasons may be, for example, tool wear or the reduction of the concentration in a solution. A progressive change in the expected value or the standard deviation of such a process is considered as systematic and not as a random cause. These are then the results of sampling, not simple random samples from the same population.

[ISO 21747, 3.1.1.6]

EDITORIAL NOTE: The English original version of this term is identically defined in ISO 3534-2 (2006) and ISO 21747 (2006); DIN ISO 21747 contains the older German translation (2007); DIN ISO 3534-2 contains the newer German translation (2013) and uses only the term "process in a state of statistical control".

Stable process

Process which is only subject to **random variation causes**

*NOTE 2: A stable process behaves in general as if **samples** drawn from the process are at any time simple **random samples** from the same **population**.*

[ISO 22514-1, 3.1.21]

*EDITORIAL NOTE: The definition of the term from the original English version ISO 3534-2 (2006) was adopted in modified form in the English version ISO 22514-1 (2014); The German version DIN ISO 22514-1 (2016) uses only the term "**stable process**".*

Statistic

Completely specified function of random variables

NATIONAL FOOTNOTE: Statistics characterize properties of a frequency distribution

[ISO 3534-1, 1.8]

Target value

Preferred value or reference value of **characteristic** which is specified in a **specification**

[ISO 3534-2, 3.1.2]

(specified) **Tolerance**

Difference between **maximum value** and **minimum value** [ISO 3534-2, 3.1.6]

Tolerance interval see specification interval

Tolerance zone see specification interval

True quantity value

Quantity value, which is in accordance with the definition of a **quantity** [VIM, 2.11]

True value

Value which characterizes a **quantity** or a quantitative **characteristic**, and which is fully defined under those conditions present in the consideration of the quantity or the quantitative characteristic

> *NOTE 1: The true value of a quantity or quantitative characteristic is a theoretical concept and generally not known exactly.*

[ISO 3534-2, 3.2.5]

Upper specification limit

Limit indicating the upper limiting value [ISO 3534-2, 3.1.4]

Variation

Difference between values of a **characteristic** [ISO 22514-1, 3.1.18]

Literature

[AIAG PPAP] AIAG Core Tools, Production Part Approval Process (PPAP), 4th edition (2006)

[AIAG SPC] AIAG Core Tools, Statistical Process Control (SPC), 2nd edition (2005)

[CDQ 0301] CDQ 0301, Mananagement of Characteristics *(Corporate directive, only available to RB internally)*

[DIN 1333] DIN 1333:1992, Zahlenangaben *(in German only)*

[DIN 55350-11] DIN 55350-11:2008-05, concepts for quality management — Part 11: Supplement to DIN EN ISO 9000:2005 *(standard series DIN 55350 currently under revision)*

[DIN EN 61710] IEC 61710:2013 Power law model — Goodness-of-fit tests and estimation methods; German version EN 61710:2013

[Epps] T. W. Epps, L. B. Pulley, A test for normality based on the empirical characteristic function, Biometrika, Vol. 70, Issue 3, 1983, pp. 723–726

[Freitag] Freitag, Zeitreihenanalyse: Methoden und Verfahren, Eul-Verlag, 2003 *(in German only)*

[Hampel] Hampel, Ronchetti, Rousseeuw, Stahel, Robust Statistics. The Approach Based on Influence Functions. Wiley, New York, 1986

[Hartung] J. Hartung, Statistik, 15. Auflage, 2009, Oldenbourg Verlag München *(in German only)*

Series Quality Management in the Bosch Group, Technical Statistics

[Booklet 1] Booklet No. 1, Basic Concepts of Technical Statistics — Continuous Characteristics

[Booklet 2] Booklet No. 2, Basic Concepts of Technical Statistics— Discrete Characteristics

[Booklet 3] Booklet No. 3, Evaluation of Measurement Series

[Booklet 7] Booklet No. 7, Statistical Process Control

[Booklet 8] Booklet No. 8, Measurement Uncertainty

[Booklet 10] Booklet No. 10, Capability of Measurement and Test Processes

[ISO 1101] DIN EN ISO 1101:2017 Geometrical product specifications (GPS) — Geometrical tolerancing — Tolerances of form, orientation, location and run-out (ISO 1101:2017); German version EN ISO 1101:2017

[ISO 21747] DIN ISO 21747:2007-03, Statistical methods — Process performance and process capability statistics for continuous quality characteristics (ISO 21747:2006; *with-drawn*)

[ISO 22514-1] DIN ISO 22514-1:2016-06, Statistical methods in process management — Capability and performance — Part 1: General principles and terminology (ISO 22514-1:2014) (ISO 22514-1:2014)

[ISO 22514-2] DIN ISO 22514-2:2015-06, Statistical methods in process management — Capability and performance — Part 2: Process performance capability and process capability statistics of time-dependent process models (ISO 22514-2:2013)

[ISO 22514-3] ISO 22514-3:2008:02, Statistical methods in process management — Capability and performance, Part 3: Machine performance studies for measured data on discrete parts

[ISO 22514-5] ISO/DIS 22514-5:2015-11(E), Statistical methods in process management — Capability and performance, Part 5: Process capability estimates and performance for attributive characteristics (in preparation)

[ISO 22514-6] ISO 22514-6:2013-02, Statistical methods in process management — Capability and performance, Part 6: Process capability statistics for characteristics following a multivariate normal distribution

[ISO 3534-1] DIN ISO 3534-1:2009-10, Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms (ISO 3534-1:2006)

[ISO 3534-2] DIN ISO 3534-2:2013-12, Statistics — Vocabulary and symbols — Part 2: Applied statistics (ISO 3534-2:2006)

[ISO 5479] DIN ISO 5479 Statistical interpretation of data — Tests for departure from the normal distribution (ISO 5479:1997)

[ISO 5725-2] DIN ISO 5725-2 Accuracy of measurement methods and results — Part 2

[ISO 9000] DIN EN ISO 9000:2015-11, Quality management systems — Fundamentals and vocabulary (ISO 9000:2015)

[ISO 9001] DIN EN ISO 9001:2015-11, Quality management systems — Requirements (ISO 9001:2015)

[ISO 80000-1] DIN EN ISO 80000-1:2013, Quantities and units — Part 1: General

[Johnson] W.P. Elderton, N.L. Johnson, Systems of Frequency Curves, Cambridge University Press (1969)

[Kölling] W. Kölling, Aus der Reihe getanzt — Identifikation von Ausreißern mit dem Hampel-Test, QZ Jahrg. 46 (2001) 3, S. 315-319 *(in German only)*

[Kruskal] W. H. Kruskal; W. A: Wallis, Use of Ranks in One-Criterion Variance Analysis, Journal of the American Statistical Association, Vol. 47, No. 260. (Dec., 1952), pp. 583-621.

[Neumann] J. von Neumann et. al., The Mean Square Successive Difference, The Annals of Mathematical Statistics, 12 (1941), 153-162

[Sachs] J. Hedderich, L. Sachs, Angewandte Statistik, Springer Verlag Berlin, 16. Auflage, 2018 *(in German only)*

[Schulze] E. Dietrich, A. Schulze, Statistical Procedures for Machine and Process Qualification, Hanser-Verlag, 7. Auflage, 2010

[Swed] F. S. Swed, C. Eisenhart, Tables for Testing Randomness of Grouping in a Sequence of Alternatives, Annals of Mathematical Statistics, 1943, XIV, 66-87

[VDA-4] VDA Volume 4, Securing the quality in the process landscape, methods, economic process design and process control (2005)

[VIM] International Dictionary of Metrology (VIM), German-English version ISO/IEC Guide 99:2007, 4th Edition (2012), published by DIN German Institute for Standardization, Beuth Verlag Berlin Vienna Zurich, ISBN 978-3-410-23472-3

[Wilrich] Graf, Henning, Stange, Wilrich, Formeln und Tabellen der angewandten mathematischen Statistik, 3. Auflage (1987) *(in German only)*

Index

Robert Bosch GmbH
C/QMM Tilsch
Wiener Strasse 42 - 46
70469 Stuttgart
Germany

Telefon +49 711 811 - 0
www.bosch.com

Page intentionally left blank

Robert Bosch GmbH
C/QMM Tilsch
Wiener Strasse 42 - 46
70469 Stuttgart
Germany
Phone +49 711 811 - 0
www.bosch.com